

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
10 May 2002 (10.05.2002)

PCT

(10) International Publication Number
WO 02/36786 A2

- (51) International Patent Classification⁷: C12N 15/82, 15/11, A01H 1/00
- (74) Agents: CAWTHORN, Christian et al.; Ogilvy Renault, Suite 1600, 1981 McGill College Avenue, Montréal, Québec H3A 2Y3 (CA).
- (21) International Application Number: PCT/CA01/01532
- (22) International Filing Date: 31 October 2001 (31.10.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
60/244,214 31 October 2000 (31.10.2000) US
- (71) Applicant (for all designated States except US): MED-ICAGO INC. [CA/CA]; 1020, route de l'Église, Bureau 650, Sainte-Foy, Québec G1V 3V9 (CA).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): VEZINA, Louis-Philippe [CA/CA]; 206, Route 138, Neuville, Québec G0A 2R0 (CA). D'AOUST, Marc-André [CA/CA]; 939, avenue Manrèse, #2, Québec, Québec G1S 2W9 (CA). ARCAND, François [CA/CA]; 1513, Pamphile-Lemay, Cap-Rouge, Québec G1Y 3C3 (CA). BILODEAU, Pierre [CA/CA]; 3477, Carré Rochon, Ste-Foy, Québec G1X 2C1 (CA).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:**
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/36786 A2

(54) Title: METHOD OF SELECTING PLANT PROMOTERS TO CONTROL TRANSGENE EXPRESSION

(57) Abstract: The present invention relates to a method for the identification and cloning of promoters that are useful in regulation of gene expression under different environmental conditions, such as in cultured transformed cells or in transgenic plants. A promoter that is a nucleic acid region located upstream of the 5' end of a plant DNA structural coding sequence that is transcribed at desired and/or modulated levels in plant tissues. The promoter regions are capable of conferring high levels of transcription in leaf tissue and in developing seed tissues when used as a promoter for a heterologous coding sequence in a chimeric gene. The promoter and any chimeric gene in which it may be used can be used to obtain transformed plant cells and plants. Chimeric genes including the isolated promoter region, transformed plants containing the isolated promoter region, transformed plant cells and seeds are also disclosed.

METHOD OF SELECTING PLANT PROMOTERS
TO CONTROL TRANSGENE EXPRESSION

BACKGROUND OF THE INVENTION

5 (a) Field of the Invention

The invention relates to plant genetic engineering and more specifically to novel methods of identifying expression regulatory sequences, including promoters and selective gene expression elements in
10 plants. The identified expression regulatory sequences are capable of conferring desired levels of transcription of heterologous genes in cells of different tissues or in *in vitro* culture. Also, novel chimeric genes selectively expressed in cells of
15 different living tissues or in *in vitro* culture, and transformed plant cells and plants containing the chimeric genes are produced.

(b) Description of Prior Art

In several cases, limitations to the
20 application of the recombinant technology has come from the inability of transgenic organisms to accumulate adequate amounts of the recombinant product, as a result of low transcription rates, improper splicing of the messenger, instability of the foreign mRNA, low
25 translation rates, hyper-susceptibility of the recombinant protein to the action of endogenous proteases or hyper-susceptibility of the recombinant organism to the foreign protein which result in improper and limited growth or in the worst cases, in
30 strong deleterious effects to the host organism. Thus, depending on the characteristics of each transgene to be expressed, different types of promoters may be

required. It is of outmost importance to have access to the appropriate promoter to control the expression of foreign genes considering various outcomes of this process: a) strategic outcome: appropriate promoters
5 are rare, and often the known one have already been patented, b) functional outcome: each protein to be express has it own requirement, so that its optimal expression and accumulation in the host cell can only be achieved with the appropriate promoter, c)
10 environmental outcome: most promoters used today in transgenic plants originated from virus and bacteria. The promoter machine aims at developing novel promoters from alfalfa using high-throughput system.

At least two key components are required to
15 stably engineer a desired trait, or control of such a trait, into a multicellular organism. The first key component comprises identifying and isolating the gene(s) which either encode(s) or regulate(s) a particular trait. The second component comprises
20 identifying and isolating the genetic element(s) essential for the expression and/or selective control of the newly isolated gene(s) so that the multicellular organism, such as a plant, will manifest the desired trait and, ideally, manifest the trait in a controlled
25 or controllable manner. This second component, which controls or regulates gene expression, typically comprises transcription control elements known as promoters. Although a generic class of promoters which drive the expression of heterologous genes in plants
30 have been identified, a broad variety of promoters active in specific target tissues or eukaryotic cells remain to be described.

- 3 -

Many systems have been used to isolate genes and their promoters located upstream of the transcription start site of a gene. The techniques can roughly be divided in three categories, namely (1) where the aim is to isolate genomic DNA fragments containing promoter activity randomly by so-called promoter probe vector systems, (2) where the aim is to isolate a gene from a genomic bank (library) and isolation of the corresponding promoter follows therefrom, and (3) where the aim is to isolate a genome fragments by PCR amplification using a known primer and a prime designed to hybridize with an adapter, a technique usually named genome walking.

In promoter probe vector systems, genomic DNA fragments are randomly cloned in front of the coding sequence of a reporter gene that is expressed only when the cloned fragment contains promoter activity. Promoter probe vectors have been designed for cloning of promoters in *E. coli* (An, G. et al., *J. Bact.* 140:400-407 (1979)) and other bacterial hosts (Band, L. et al., *Gene* 26:313-315 (1983); Achen, M. G., *Gene* 45:45-49 (1986)), yeast (Goodey, A. R. et al., *Mol. Gen. Genet.* 204:505-511 (1986)) and mammalian cells (Pater, M. M. et al., *J. Mol. App. Gen.* 2:363-371 (1984)). It is known in the art that, for example, promoters of different organisms fail to work in *E. coli* and yeast (e.g. Penttila, M. E. et al., *Mol. Gen. Genet.* 194:494-499 (1984)). Therefore, these microorganisms cannot be used as hosts to isolate such promoters, and most probably promoters from a multiple of other higher organisms.

Known genes can be isolated from either a cDNA or chromosomal gene bank (library) using hybridization as a detection method. Such hybridization may be with a corresponding, homologous gene from another organism or
5 with a probe designed on the basis of expected similarities in amino acid sequence. If amino acid sequence is available for the corresponding protein, an oligonucleotide can also be designed which can be used in hybridization for isolation of the gene. If the gene
10 is cloned into an expression library, the expression product of gene can be also detected from such expression bank by using specific antibodies or an activity test.

However, a major concern is how to isolate
15 specific genes that have the desired promoter properties, for example promoters which would allow for most highly expression in selected conditions, as in the cases of different desired combinations of promoters and encoding sequences into a DNA expression
20 vector. There is little information available in the literature to indicate which genes are the most highly expressed in several organisms. In addition, it would be useful to have a method for isolating promoters with a mean which is not dependent on specific mRNA relative
25 abundance.

High technologies in genetic, bioinformatic and robotic enable life science scientists to study biological processes which are very complex. The same technologies and the knowledge gain through them can
30 subsequently be applied to meet specific needs in the area of health, agriculture and environment. Genetic mapping and DNA sequencing of complete genome is one of the most outstanding demonstration of the power of the

high technologies which was possible through the automation of DNA sequencing and DNA fragment isolation protocols. Several prokaryote and eukaryote genomes have already been completely sequenced including
5 bacteria (*E. coli*), yeast (*S. cerevisiae*), worm (*C. elegans*), and plant (*A. thaliana*). The scientific community is now well on its way to obtain the complete genome sequence of several other organisms including the human genome. The next step of these
10 megasequencing project is the identification of putative open reading frames (ORFs), which are the sequences that will be translated in amino acid sequences (proteins). But more importantly, this will lead to the identification of the functions of each of
15 these proteins within their immediate cellular environment and within the whole organism. This knowledge of the proteome (the entire protein population in a given environment) will allow the understanding of all the biological processes in this
20 environment. The present invention provides specific strategies (promoter machine and protein machine) to acquire specific molecular tools. This invention also identifies different ways by which these tools can be applied to enable further development in genomic and
25 proteomic research.

The activation of DNA promoter is a very complex process. The expression of the genes occurs sequentially, probably as the result of a "cascade" mechanism of transcriptional regulation. Thus, an
30 immediate-early gene may be expressed immediately after activation, in the absence of other functions, and one or more of the resulting gene products induces transcription of the delayed-early genes. Some delayed-

early gene products, in turn, induce transcription of late genes, and finally, the very late genes are expressed under the control of previously expressed gene products from one or more of the earlier classes.

5 Activation of a promoter is influenced by several factors, even by the gene itself to which it is linked. Production efficiency of recombinant polypeptides in transgenic cells and organisms is often dependent on these facts, putting out that combination of promoters

10 and genes of interest can almost always be both quantitatively and qualitatively improved.

Thus, there is a great need for a system that expresses foreign gene products in a continuous and permanent manner such that the cell is still capable of

15 processing products efficiently. Described herein is such a system, as well as an improved and novel vector for gene expression employing selected combination of promoter-gene vectors.

Also, it would be highly desirable to be

20 provided with selective promoters and with a method of isolating and characterizing a large number of promoters, as well as an issuing method of application-customized scale production system using selected promoters in genetically transformed organisms and

25 microorganisms.

SUMMARY OF THE INVENTION

Access to promoters would enable the genetic engineering of tissues or eukaryotic cells from

30 commercially important organisms such as agricultural animal and plants, and microorganisms. Screening of DNA libraries was undertaken as a method for the

identification promoters from eukaryotic organisms and microorganisms. Such sequences can be identified, and the promoters and their associated structural genes sequenced. Expression of genes encoding for polypeptides and/or RNA in alfalfa plants is used as an assay of the tissue specificity and other characterizations of the isolated promoters and DNA vectors.

One object of the present invention is to provide plant tissue selected expression regulatory sequences and DNA vectors, containing the selected expression regulatory sequence and gene encoding a desired protein, adapted for specific applications.

Another object of the present invention is to provide a method of producing adapted DNA vector for expression and/or production of recombinant polypeptides and/or RNA comprising the steps of:

- a) isolating mRNA from cells;
- b) preparing a cDNA library from the mRNA;
- 20 c) producing at least one oligonucleotide primer from cDNAs of the cDNA library of step b), the oligonucleotide primer allowing amplification of promoter and/or signal peptide upstream of the cDNAs;
- 25 d) performing amplification of at least one expression regulatory sequence upstream or downstream of a cDNA, a genomic DNA sequence with the oligonucleotide primer of step c) on a genomic DNA sample;
- 30 e) linking the amplified sequence of step d) to a gene encoding for a directly or indirectly

detectable polypeptide and/or RNA to form a DNA expression vector for expression of the detectable polypeptide; and

- 5 f) selecting a DNA expression vector or expression regulatory sequence of step e) by measuring levels of expression of the detectable polypeptide and/or RNA under conditions allowing activation of the promoter and expression of the detectable
10 polypeptide and/or RNA.

In accordance with the present invention there is provided a method that use mRNA from different cell types, such as plant, animal, mammal, or cells to produced cDNAs. Also, such cDNAs can be used to
15 produce recombinant polypeptide and/or RNA in genetically transformed cells. An oligonucleotide primer sequence can be also determined starting from a cDNA or any DNA fragment available in a data bank, or even from a synthetic DAN fragment.

20 In accordance with the present invention there is provided a method, wherein the polypeptide and/or RNA origin from the group consisting of pharmaceutical, agronomic, environmental, industrial, nutraceutical, cosmeceutical polypeptide, gene product markers, fusion
25 protein, green fluorescent protein, and β -glucuronidase.

Also, the method of the invention may be performed in vitro in transitory transfected cells or stably genetically transformed cells, as well as in
30 vivo, in a seed or a growing organism.

The detection and measurement of polypeptide and/or RNA may be indirectly detected by using

antibodies, Western blot, Northern blot, *In situ* hybridization, colorimetry, optical densitometry, spectrophotometry, and/or migrating gels. The polypeptide may comprise a tag, self cleavable in
5 certain cases, to be directly detected or for purification of the polypeptide and/or RNA.

In accordance with the present invention there is provided an expression regulatory sequence, which is natively located upstream or downstream of a gene
10 encoding a polypeptide and/or RNA and controls the expression of a gene encoding a polypeptide and/or RNA.

Another object of the invention is to provide with a transgenic plant regenerated from stably genetically transformed cells with selected
15 combinations of an expression regulatory element according to the present invention and at least one gene, or a DNA vector which may be a plasmid vector or a viral vector.

In accordance with the present invention there
20 is provided plant cells and transgenic plants transformed with DNA vectors of the in the present invention.

Another object of the present invention is to provide a method of isolating and characterizing an
25 expression regulatory sequence for expression of a recombinant polypeptide and/or RNA comprising the steps of:

a) producing at least one oligonucleotide primer from a cDNA, genomic DNA fragment or
30 synthetic DNA sequence, the oligonucleotide primer allowing amplification of a genomic sequence upstream or downstream of a genomic

complementary site of the oligonucleotide primer;

- 5 b) performing amplification of the genomic sequence upstream or downstream of the genomic complementary site of the oligonucleotide primer a) on a genomic DNA sample;
- 10 c) linking an amplified sequence obtained from the amplification of step b) to a gene encoding for a directly or indirectly detectable polypeptide and/or RNA to form a DNA expression vector for expression of the detectable polypeptide and/or RNA; and
- 15 d) selecting at least one expression regulatory sequence from the vector of step c) by measuring levels of expression of the detectable polypeptide and/or RNA under a condition allowing activation of the expression regulatory sequence and expression
- 20 of said detectable polypeptide and/or RNA.

Also, another object is to provide a method of producing an adapted DNA vector for expression of recombinant polypeptides and/or RNA comprising the steps of:

- 25 a) producing at least one oligonucleotide primer from a cDNA, a genomic DNA fragment or a synthetic DNA sequence, the oligonucleotide primer allowing amplification of a genomic sequence upstream or downstream of a genomic
- 30 complementary site of the oligonucleotide primer;

5 b) performing the amplification of the at least one genomic sequence upstream or downstream of the genomic complementary site with the oligonucleotide primer of step a) on a genomic DNA sample;

10 c) linking an amplified sequence obtained from the amplification of step b) to a gene encoding for a directly or indirectly detectable polypeptide and/or RNA to form a DNA expression vector for expression of the detectable polypeptide and/or RNA; and

d) selecting a DNA expression vector of step c) by measuring the level of expression of said detectable polypeptide and/or RNA.

15 For the purpose of the present invention the following terms are defined below.

The term "polypeptide" as used herein, refers to any amino acid sequence, oligopeptide, peptide, or protein sequence, or a fragment of any of these, and to
20 naturally occurring or synthetic molecules. Where "polypeptide" is recited herein to refer to a polypeptide sequence of a naturally occurring protein molecule, "polypeptide" and like terms are not meant to limit the amino acid sequence to the complete native
25 amino acid sequence associated with the recited protein molecule. The "polypeptide" may be endogenous, exogenous, naturally occurring or recombinant.

The term "complementary" as used herein is intended to mean a recognition DNA sequence that is
30 complementary to another sequence, such a primer can recognize and anneal with complementary site or sequence in a genomic DNA sample. Complementary

characteristic partial, since an oligonucleotide primer can anneal on a partial distance to a recognition site in a DNA sample.

The expressions "coding sequence" and
5 "structural sequence" refer to the region of continuous sequential DNA triplets encoding a protein, polypeptide, or peptide sequence.

The term "linked" meaning also "coupled", refers to a promoter or promoter region and a coding or
10 structural sequence in such an orientation and distance that transcription of the coding or structural sequence may be directed by the promoter or promoter region.

The term "expression" as used herein means the transcription of a gene to produce the corresponding
15 mRNA and translation of this mRNA to produce the corresponding gene product, such as a peptide, polypeptide, or protein.

The term "gene" refers to chromosomal DNA, plasmid DNA, cDNA, synthetic DNA, or other DNA that
20 encodes a peptide, polypeptide, protein, or RNA molecule, and regions flanking the coding sequence involved in the regulation of expression.

"Overexpression" refers to the expression of a polypeptide or protein encoded by a DNA introduced into
25 a host cell, wherein the polypeptide or protein and/or RNA is either not normally present in the host cell, or wherein the polypeptide or protein is present in the host cell at a higher level than that normally expressed from the endogenous gene encoding the
30 polypeptide or protein.

- 13 -

The expression "expression regulatory sequence" as used herein refers to a promoter, a promoter region a transcription regulatory sequence, a DNA sequence usually found upstream (5') or downstream (3') to a coding sequence, involved in the control of expression of the coding sequence by controlling production of messenger RNA (mRNA) by providing the complementary site for RNA polymerase and/or other factors necessary for initiation of transcription at the correct site. As contemplated herein, an expression regulatory sequence includes variations of promoters derived by means of ligation to various regulatory sequences, random or controlled mutagenesis, and addition or duplication of enhancer sequences. The expression regulatory sequence disclosed herein, and biologically functional equivalents thereof, are responsible for driving the transcription and translation of nucleic acid sequences under their control when introduced into a host as part of a suitable recombinant vector, as demonstrated by its ability to produce mRNA. An expression regulatory sequence may be also a 3' regulatory sequence, such as, but not limited to, 3' UTR element, acting as a stabilizing agent of during the processing of the RNAs in a cell. An expression regulatory sequence can be a regulatory element.

The expression "regulatory element" as used herein refers to a DNA sequence that can increase or decrease the amount of product produced from another DNA sequence. The regulatory element can cause the constitutive production of the product (e.g., the product can be expressed constantly). Alternatively, the regulatory element can enhance or diminish the production of a recombinant product in an inducible

fashion (e.g., the product can be expressed in response to a specific signal). The regulatory element can be regulated, for example, by nutrition, by light, or by adding a substance to the transgenic organism's system.

5 The terms "recombinant DNA construct" or "recombinant vector" or "DNA vector" as used herein mean any agent such as a plasmid, cosmid, virus, autonomously replicating sequence, phage, or linear or circular single-stranded or double-stranded DNA or RNA
10 nucleotide sequence, derived from any source, capable of genomic integration or autonomous replication, comprising a DNA molecule in which one or more DNA sequences have been linked in a functionally operative manner. Such recombinant DNA constructs or vectors are
15 capable of introducing a 5' regulatory sequence or promoter region and a DNA sequence for a selected gene product into a cell in such a manner that the DNA sequence is transcribed into a functional mRNA which is translated and therefore expressed. Recombinant DNA
20 constructs or recombinant vectors may be engineered to express a large number of polypeptides of interest.

"Transformation" refers to the introduction of DNA into a recipient host or hosts. "Host" or "hosts" refers to bacteria, entire plants, plantlets, or plant
25 parts such as plant cells, protoplasts, calli, roots, tubers, propagules, seeds, seedlings, pollen, any other plant tissues, and other eukaryotic organisms and microorganisms.

30 BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates a schematic representation of the strategies involved in the Promoter and Proteins

Machines. Each box corresponds to a specific task of the invention and the arrows indicate the links between each task;

Fig. 2 illustrates according to one embodiment
5 of the present invention a schematic representation of genomic walking;

Fig. 3 illustrates Inducibility and expression level of the GUS gene in tobacco leaves using the Nitrite Reductase upstream and downstream sequences;

10 Fig. 4 illustrates the GUS expression level in transgenic tobacco leaves under the control of alfalfa Plastocyanin upstream and downstream sequences; and

Fig. 5 illustrates the detection of protein X in alfalfa cell cultures by Western.

15

DETAILED DESCRIPTION OF THE INVENTION

In accordance with the present invention, there is provided a method of isolating and characterizing a large number of known and unknown promoters from cells
20 in a same assay. Such isolated promoters are then operably linked to a gene, or cDNA encoding for a protein of interest, to form a DNA expression vector for which the expression efficiency is assessed in cultured cells or whole organisms.

25 The present invention, in at least one of its aspects, relates to one or more DNA sequences that can be used as promoters for expressing endogenous or foreign genes in plant cells and/or plants, most particularly in alfalfa.

The DNA sequences of the present invention include at least an effective part of a sequence present in a vector obtained from a genomic library of alfalfa. By term "effective part" is meant a part of
5 the indicated DNA sequence that, when fused to a particular gene and introduced into a plan cell, causes expression of the gene at a level higher than is possible in the absence of such part of the indicated DNA sequence.

10 For the purpose of the present invention, it is not critical which transformation technique is used, provided it achieves an acceptable level of gene transfer in cells or an organism.

To be able to develop versatile systems for
15 protein production from transformed plants, especially when plants are grown, a method has been developed for the isolation of previously unknown alfalfa genes which are highly expressed, and their promoters. The method of the invention can require, but is not limited to,
20 the use of only one cDNA population of probes.

It is to be understood that the method of the invention, for certain applications, is useful for the identification of promoter sequences that are active under any desired environmental condition to which a
25 cell may be exposed, and not just to the exemplified isolation of promoters that are capable of expression in specific conditions. By "environmental condition" is meant the presence of a physical or chemical agent, such agent being present in the cellular and plant
30 environments, either extracellularly or intracellularly. Physical agent would include, for example, certain growth temperatures, especially a high

or low temperature. Chemical agents would include any compound or mixtures including carbon growth substrates, drugs, atmospheric gases, etc.

Also, once the genetic material (DNA) of a
5 given organism has been completely sequenced, it can be used to isolate and identify the proteins that are encoded by the DNA sequences. In order to study and understand the specific function of these proteins, they must be either expressed in heterologous system or
10 extracted from their host. Considering the enormous number of proteins, this is the limiting step. Few expression systems are currently available to enable the expression of these unknown proteins and they all have their own limitations. The present invention
15 provides an additional expression system, the protein machine, that allow the use of selected promoters developed through the promoter machine to rapidly produce small quantities of recombinant proteins in plant cells. The protein machine uses current
20 protocols in cell culture, plant cell transformation, and recombinant protein purification in a high-throughput system

According to the method of the invention, the organism may be first grown under the desired growth
25 condition, such as in *in vitro* culture or *in vivo*. Total mRNA is then extracted from the organism and preferably purified through at least a polyA+ enrichment of the mRNA from the total RNA population. A cDNA bank, or cDNA library is made from this total mRNA
30 population using reverse transcriptase and the cDNA population cloned into any appropriate vector, such as the commercially available lambda-ZAP vector system (Stratagene). When using the lambda-ZAP vector system,

- 18 -

or any lambda vector system, the cDNA is packaged such that it is suitable for infection of any *E. coli* strain susceptible to lambda bacteriophage infection.

The cDNA bank is transferred by standard colony
5 hybridization techniques onto nitrocellulose filters for screening. The bank is plated and plaque lifts are taken onto nitrocellulose. The bank is screened with a population of labeled cDNAs that had been synthesized against the same RNA population, from which the cloned
10 cDNA bank was constructed, using stringent hybridization conditions. This results in clones hybridizing with varying intensity and the ones showing the strongest signals are picked. Genes that are most strongly expressed in the original population comprise
15 the majority of the total mRNA pool and thus give a strong signal in this selection.

The inserts in clones with the signals are sequenced from the 3' end of the insert using any standard DNA sequencing technique as known in the art.
20 This provides a first identification of each clone and allows the exclusion of identical clones. The frequency with which each desired clone is represented in the cDNA lambda-bank is determined by hybridizing the bank against a clone-specific PCR probe. The desired clones
25 may be those which, in addition to having the strongest signals as above, are also represented at the highest frequencies in the cDNA bank, since this implies that the abundance of the mRNA in the population was relatively high and thus that the promoter for that
30 gene may be highly active under the growth conditions. It is very important also to note that mRNA abundance may be dependent on the stability to the mRNA itself. Thus, the relevance of this approach and any clone

identified therefrom can be double-checked: the intensity of the hybridization signal of a specific clone should correlate positively with the frequency with which that clone is found in the cDNA library. The
5 inserts of the clones selected in this manner, such inserts corresponding to the cDNA sequences, may be used as probes, routinely named EST, to isolate the corresponding genes and/or their promoters from a genomic bank, such as one cloned into lambda as above.

10 The method of the invention is not limited to plants, but would be useful for cloning genes from any host, or from a specific tissue with such host, from which a cDNA library may be constructed, including, prokaryote (bacterial) hosts, and any eukaryotic host
15 plants, mammals, insects, yeast, and any cultured cell populations.

In a preferred embodiment of the invention, isolation of promoters, combination with desired encoding gene, and selection of optimum DNA vector thus
20 form including these sequences, may be performed in a high throughput automated system.

The indicated fragments of the present invention can be fused to foreign genes of diverse origins and incorporated into vectors designed for genetic
25 transformation of plants and then used in standard genetic engineering techniques. For example, an isolated fragment according to the present invention may be linked to a target gene that encodes a functional protein, reporter polypeptide or RNA. The
30 gene linked to the promoter fragment may be an endogenous gene (or cDNA fragment) or a foreign gene (or cDNA fragment) isolated from any other source.

The emerging industry of molecular farming (production of recombinant molecules in animals or crops) is one of the most promising industry of the coming century. It is of particular embodiment of the present invention to provide safe and renewable molecule factories for the industry. Among the applications that are currently developed are the production of low-cost monoclonal antibodies for therapeutic and diagnostic uses, the production of unlimited amounts of hormones, cytokines and other bio-active molecules for the treatment of chronicle or lethal diseases, the production of bio-safe substitutes for various blood components, the production of unlimited amounts of processing enzymes for the food and pulp industry, the production of low-cost enzymes for waste treatments, and the production of safe bio-active molecules for the cosmetic industry.

Of particular embodiments, the method of the present invention can be used also for the identification and isolation of analogous promoters, signal peptides and structural genes in several species of multicellular and unicellular organisms.

Another important aspect of the invention is the improvement of the expression efficiency in transgenic plants containing adapted DNA vector as described above, in terms that it may be more controllable quantitatively and qualitatively in producing recombinant proteins, polypeptides and RNAs.

The subject promoter sequences find a wide variety of applications. In one embodiment, the subject sequences are used to regulate the synthesis of polypeptides which in turn provide a number of

applications, including use in proteomic microarrays, models for rational drug design, immunogens for antibody elicitation, etc.

As yet in a preferred embodiment, the present invention can be preformed in an automated high throughput system. Screening of most efficient combinations of promoter-gene may be rapidly carried out, and production of a large number of clones allowing availability of as many choices of polypeptides for proteomic protocols and drug targeting. Therefore, the invention may be used also as a high throughput identification system of candidate therapeutic targets.

In the most preferred embodiment, the method of the invention provides capacity to produce large quantities of stably-transformed alfalfa cell lines constructed to express a heterologous DNA of interest under the control of different promoters or combination of promoters and other regulatory sequences. In its most preferred embodiment, a combination of promoter-gene, a DNA vector, therefore selected allows for preparation of genetically transformed alfalfa cell lines and plants, performing themselves expression at a desired level of polypeptides for a specific applications. Polypeptides can be produced on an application-specific-scale basis or on a large-scale basis.

Important embodiments of the invention are; high throughput promoter machine able to perform a series of automated manipulations aiming at isolating interesting DNA fragment that posses promoter activities with known gene expression patterns; cDNA libraries from various

alfalfa tissues (leaves, cell cultures); adapted genomic libraries from alfalfa; nucleotide sequence database of genes expressed in alfalfa leaves and cell cultures; alfalfa DNA chips and DNA microarray
5 information; database of oligonucleotides specific to given EST sequences; a database of genomic DNA sequences native to alfalfa which are involved in gene regulation; a database of cryptic DNA sequences active in the regulation of gene expression in alfalfa leaves
10 and cell culture; a database of DNA sequences representing the transcriptional machinery of alfalfa leaves and cell culture; a database of synthetic oligonucleotides responsible for various gene expression patterns in alfalfa; a database of plant
15 promoters responsible for specific gene expression patterns in alfalfa; a high throughput protein machine able to perform a series of automated manipulation aiming at producing small quantities of proteins from various sources; small quantities of proteins (mg) from
20 various sources to be tested for bioactivity; and small quantities of proteins (g) from various sources to be used in pre-clinical trials and different types of study.

The present invention will be more readily understood by referring to the following examples which
25 are given to illustrate the invention rather than to limit its scope.

EXAMPLE I

30

Promoter machine

Production of cDNA libraries from specific tissue

The first step is to produce cDNA libraries

which serve as starting material for this invention. The quality of the libraries is very important as they must represent the complete mRNA population from their tissue of origin and they must contain full-length cDNA clones both at the 5' and 3' end of the mRNA molecule. Therefore, these libraries are made manually using commercially available kit. They can be either phagemid or plasmid libraries. One of the major applications of this invention is in molecular farming using cell culture and/or whole plants of alfalfa (*Medicago sativa*). Therefore, the tissue from which the cDNA libraries can be derived are leaves and alfalfa cell culture. Although plant cell cultures are sometimes derived from leaf cells, it is likely that cell culture will not express the same genes as leaf cells.

The production of cDNA libraries and subsequent sequencing of the EST clones is the major point of entry of the promoter machine high throughput system. However, it may not be the only one. Novel promoters can also be generated using the DNA sequences available in our laboratories. So far, this experiment has produced more than 137 588 ESTs which are publicly available. Primers can be derived from these EST sequences and used directly for the genome-walking step without having to sequence any EST from alfalfa cDNAs. Similarly, synthetic promoters can be constructed using random oligonucleotides hooked to a minimal promoter. This example is outlined in Example H.

30 EST sequencing

Once made, the cDNA libraries are cultivated on petri dish to generate a number of independent clones.

For the high throughput system, each of these independent clones is selected either manually or through an automated process to allow for its amplification, storage and isolation of the
5 corresponding plasmid DNA. These procedures can be done using standard protocols such as the Biomek 2000™ double stranded DNA isolation of DNA sequence templates as used in different laboratories. Following DNA sequencing, EST sequences are automatically loaded into
10 a database for further analysis. In most DNA sequencing protocols, only the 5' sequences of ESTs are obtained which most likely containing the major part of the coding sequence of the ESTs. For comparative studies, it is known that the non-coding sequences
15 (promoters and 3' non-coding) are not as conserved as the coding sequences (ORFs). Therefore, sequencing of 5' portions of ESTs allow for better comparisons of genes between species. In addition, to generate the necessary information to design PCR primers for the
20 genome walking protocol, 5' sequences of the

ESTs must be obtained. Since regulatory sequences are found upstream or downstream of the initiating ATG, the 5' sequences also allow the identification of the appropriate sequence that can be
25 used as DNA template to design the PCR primers (oligonucleotides) used for amplification of corresponding regulatory sequences. The 5' sequences may also provide valuable information as to whether the ESTs are full length, if the EST contain signal peptide
30 (transit peptide, cleavage site, etc) and/or if the EST are homologous to other sequences previously identified in the same species and/or in different species.

- 25 -

In addition, we can also sequence the 3' end of each EST. This provides additional information used to identify the nature and the potential role of the EST. For example, in the case of isoenzymes, two genes of the same gene family may have almost identical coding sequences but be expressed very differently. In this example, the 5' sequences (coding) of the corresponding ESTs are identical but the 3' sequences might be very different. Sequencing the 5' region only may reflect a gene duplication while sequencing the 3' region which are likely to be different can detect the presence of two different genes. In another situation where two clones originating from the same gene are of different length, the 5' sequence of the corresponding ESTs shows two different sequences but the 3' sequences is identical.

Production of adapted genomic libraries

This section enables the production of one of the two major components of the Genome walking strategy. Adapted genomic libraries are produced from the selected organism (for example alfalfa) to serve as a template DNA for specific PCR amplifications. These genomic libraries are produced manually considering their quantitative and qualitative importance. The adapted genomic libraries are constructed on the same principle as a convention phagemid library using standard protocols and genomic DNA digested with specific DNA restriction enzymes. However, one of the differences is that known DNA sequences are placed at each end of the resulting DNA fragment in order to use these known sequences at a later stage as primer for PCR amplification. In order to increase the probability of amplifying a specific PCR fragment for a

given EST sequence during genome walking, several different genomic libraries can be constructed with the same known sequences at each end but using different restriction enzymes to digest the genomic DNA. Once
5 constructed, the adapted genomic libraries are amplified and the DNA can be extracted and used as template DNA for the PCR amplifications.

The adapted genomic libraries can also be used in a sequencing project to obtain additional DNA
10 sequence information from a given organism (for example alfalfa). Sequencing of genomic clones reveals different type of information than sequencing EST clones. The genomic clones contain non-coding regions (promoter, terminator, introns, 5' leaders, spacers,
15 repeated regions, pseudogenes, etc) while EST clones contain principally coding regions and open reading frames. Sequences of non-coding regions are valuable tools for comparative studies between members of the same species and/or members of different species.

20 PCR primer design

A second component of the genome walking strategy is a pair of oligonucleotides (proximal and distal) to be used as primers in nested PCR amplification on the genomic DNA extracted from the
25 adapted genomic libraries. The aim of these amplifications is to isolate and clone the 5' regulatory sequences located upstream of the proximal part of each EST in the genome of the corresponding organism. The oligonucleotides are derived from the 5'
30 sequence of the EST in the reverse orientation from the reading frame. For example considering that the reading frame is in 5' to 3' orientation, the two

- 27 -

oligonucleotides (proximal and distal) would be made from the 3' to 5' orientation of the same reading frame. In addition, they must be separated by a reasonable length of DNA so that the nested PCR
5 amplification can be performed and the corresponding PCR fragments would be different enough in size to be differentiated by electrophoresis on a 2% agarose gel. The design of these oligonucleotides from the EST sequences is a tedious task considering that thousands
10 of sequences are generated. Therefore, appropriate software is used to perform this task and the oligonucleotide sequences selected are fed directly into an oligo synthesizer that produce the primers. Alternatively, this task may alternatively be done
15 manually.

Genome walking

This section is one of the steps of the present invention since its application in a high throughput system has not yet been attempted. Fig. 2 represents a
20 summary of the steps involved. A first PCR amplification is performed using the distal primer derived from the sequence of the EST and another primer derived from the known sequences located at each end of the genomic clones in the adapted genomic libraries. A
25 second PCR amplification is performed on the first PCR reaction mixture using the proximal primer and the known adapter primer. The following step is to confirm the amplification of a specific DNA fragment in the second PCR reaction by gel electrophoresis. The
30 visualized amplification product can be cloned into a PCR fragment-cloning vector. This DNA fragment corresponding to the promoter region of a given EST clones can be sequenced using primers corresponding to

the vector sequences flanking the inserted genomic fragment. The presence of identical DNA sequences between the EST sequence and the corresponding promoter sequence (PCR amplified DNA fragment) would confirm
5 that the amplified promoter is the native sequence controlling the gene expression of a given EST.

This step is considered as the time-regulating step of the entire system of the promoter machine. This is due to the fact that this step contains many
10 subtasks to be performed including two PCR amplifications, the detection of a PCR fragment by gel electrophoresis or other means, and the subcloning and DNA sequencing of this same PCR fragment. Furthermore, the sequence generated have to be analyzed in order to
15 confirm if the corresponding DNA fragment is the regulatory region (promoter) associated with the EST sequences used to generated the PCR primers (proximal and distal). By the end of this section, a large number of promoter sequences have been identified and
20 analyzed. These promoters are flanked at their 5' end by known sequences corresponding to the adapter used to PCR amplified them.

Ligation promoter-reporter gene

PCR fragment generated during the genome-walking
25 step must be isolated in large enough quantities to be able to ligate them to the appropriate cloning vector. These cloning vectors are prepared in advance following construction, amplification and linearisation with the appropriate restriction enzyme. They contain a
30 reporter gene to be fused to the promoter fragment isolated by the genome walking protocol. The reporter gene must be easily detectable by common method. For

example, the B-glucuronidase (GUS) gene and the Green Fluorescent Protein (GFP) gene can be used. Their gene products can be easily detected by spectrophotometric or fluorometric analysis.

5 The promoter-reporter gene fusion can be transcriptional and/or translational fusion. For transcriptional fusion, only the regulatory sequence must be ligated while for translational fusion, part of the coding sequence can be ligated but it must be in
10 the appropriate reading frame in order to be functional. For transcriptional fusion, the sequence of the promoter region is analyzed and the initiating ATG is identified. Then a new oligonucleotide containing the ATG region is generated in the
15 orientation 3' to 5' compare to the normal reading frame. The promoter fragment is amplified again by PCR using the new oligonucleotide and the genomic primer derived from the known sequences flanking the genomic DNA in the adapted genomic libraries. At the same
20 time, the reporter gene is also amplified by PCR using two specific primers, one of which is derived from the initiating ATG of the reporter gene but also contains complementary sequences to the new oligonucleotide used to amplified the expression regulatory sequence
25 fragment. To make the transcriptional fusion, the two resulting PCR fragment, promoter of interest and reporter gene, are placed together and used in a third PCR amplification using the primer located at the 5' end of the promoter fragment and the primer located at
30 the 3' end of the reporter gene. The resulting PCR fragment should contain the transcriptional fusion between the promoter of interest and the reporter gene and can then be inserted into a cloning vector for

further experiments. This type of ligation is generally well known by those skilled in the art.

For translation fusion between promoter and reporter gene, each PCR fragment corresponding to a promoter region is ligated into three different cloning vectors. Each of these three cloning vectors represent one potential reading frame so that one out the three ligation events should contain the translational fusion between the promoter of interest and the reporter gene. The two other vectors containing the fusion not in frame should not be detected at the gene expression analysis step since the reporter gene should not be translated correctly. In the event that the PCR fragment generated only contain the promoter region of the gene, the three translational fusion should give the same expression pattern. The translational fusion has the additional advantage that it may identify other regulatory sequences apart from the promoter itself. Regulatory sequences have been found before in introns, 5' leader sequences and 3' leader sequences. In addition, the translated part of the gene of interest might contain signal peptide that would target the accumulation of the reporter gene into a specific cellular localization. Following histochemical localization of the product of the reporter gene, it might help in the identification of novel signal sequences.

Cell transfection

Several techniques are available perform integration of DNA into plant cells including, but not limited to, *Agrobacterium*-mediated transformation, silicon carbide whiskers, biolistic protocol (gene

gun), and direct transfer method using PEG, electroporation and/or cationic polymers. Under certain considerations, it may not be realistic to undertake the task of producing mature transgenic plants with each individual construct generated. Knowing that particular tissue type can be transformed with specific transformation techniques, the tissue type must be identified before we consider the technique to use. In the event that leaves are used as plant material to transform, transfer of plasmid DNA by biolistic would be an appropriate method of transformation. On the other hand, if plant cell culture and/or protoplasts are used as starting material, direct transfer methods such as PEG, electroporation and cationic polymers can be used.

Promoter activity analysis

This section aims at determining the expression patterns controlled by the promoters of interest. Following transformation of the plant cells with the cloning vectors containing the promoter of interest fused to the reporter gene, the transformed plant cells are incubated for a period of time to allow the expression of the reporter gene under the control of the promoter of interest. Then, the same transformed plant cells are analyzed in order to quantify the activity of the promoters. When the reporter gene is the GUS gene, the transformed plant cells are put in contact with the appropriate substrate which is converted to a detectable product following conversion by the GUS gene product. This product is detectable by spectrophotometric analysis. When the reporter gene is GFP, the transformed plant cells are directly analyzed for presence of the GFP gene product by fluorometric

analysis using the appropriate wavelengths. To optimize the detection of the reporter gene product, transformed plant cells may have to be homogenized by mechanic means (PolytronTM, blender, glass beads, etc).

5 For each promoter analyzed, recorded data are compared to a negative control. For example, a negative control may be the expression of a reporter gene without any promoter fused to it. A positive result would be any promoter activity that is significantly higher than the

10 negative control. The quantification of the expression level controlled by each promoter construct should be done in triplicate to minimize the possibility of errors.

Identification of interesting promoters

15 The objective of the promoter machine is to isolate and characterize a number of promoters that drive the expression of a reporter gene within a desirable range in the desired tissue type, or based on any other criteria.

20 Protein Machine

Ligation promoter of interest:gene of interest

In contrary to the Promoter Machine system where the reporter gene (gene of interest) was inserted in the cloning vector and the promoters were ligated into

25 this vector afterward, in the Protein Machine System, it is the promoter that can be inserted in the cloning vector and it is the gene of interest that can be ligated afterward in the same vector. In the case where known proteins (genes previously isolated and

30 characterized) are used, the corresponding coding sequence of these genes are PCR amplified and inserted appropriately in the cloning vector containing the

desired promoters. Depending on the total number of promoters used and the total number of independent proteins to express this step can be either automated or manual. In addition, both transcriptional and/or
5 translational fusion can be done.

In the case of unknown proteins (a mixture of cDNA clones or EST clones obtained from a tissue-specific cDNA library or a mixture of PCR fragments obtained from a similar source) when the reading frame
10 is not known readily, a translational fusion might have to be done. To achieve this, each independent gene of interest have to be ligated to three different cloning vector each of which represent one of the three potential reading frames. This means that for each
15 promoter construct, three different cloning vectors have to be made. In addition to the importance of the reading frame used, another variable must be considered. The proteins of interest that are produced in the protein machine are extracted and purified
20 subsequently. If these proteins are unknown, this means that the cloning vector should account for specific tools to allow the purification of these unknown proteins. For example, these tools can be known antibody recognition sites, peptidic tags, his
25 tags, GST fusion, etc. These tags would allow the purification of the desired proteins through affinity chromatography techniques. Another possibility would be to do a protein fusion between the protein of interest and a protein easily detectable by
30 spectrophotometric means such as GUS and/or GFP.

Similarly to the Promoter Machine System, once the ligation into the appropriate cloning vector is completed; the resulting DNA plasmids are transformed

in bacteria for amplification. When known genes of interest are studied, the plasmid DNA from a dozen of bacterial colonies are extracted for each transformation event in order to confirm the presence
5 of the same plasmid DNA in each independent colony. Following this confirmation, the plasmid DNA is transferred in plant cell culture for expression of the gene of interest. When unknown genes of interest are studied, a great number of independent bacterial
10 colonies are selected; the plasmid DNA from each of them is extracted and sequenced. The sequence analysis should allow for the confirmation of the insertion of a unique gene of interest in the cloning vector, for the nature of the gene and the corresponding gene product,
15 and for the analysis of the reading frame in which the gene of interest has been inserted into the cloning vector.

Cell transfection

This step is performed in the same way as in the
20 Promoter Machine. The selected DNA plasmids isolated in the previous step is transferred directly in plant cell culture by the same methods described above. The resulting transformed plant cells are incubated to allow for the detection, the extraction and the
25 purification of the heterologous proteins.

Expression analysis

In the Promoter Machine System, the expression analysis was possible by the detection of the reporter gene GUS and/or GFP. In the Protein Machine system,
30 this detection is possible using the gene product itself (if known) or using specific tools (peptidic tags) fused to the gene product. Quantitative and

qualitative characterization of the produced protein is performed according to the specific characteristics of each protein.

Cell culture and protein purification

Following the analysis of the expression of each gene of interest under the control of each selected promoter of interest, the cell cultures expressing the highest level of proteins is selected. They are incubated in optimal culture conditions and proteins are then be extracted and purified in order to study them and determine their function *in vivo*. The Protein Machine system permits the expression of a great number of proteins from various sources. If greater amount of certain proteins is required for different applications (commercial or academic), the selected proteins may be produced directly in larger volumes of cell culture or in transgenic plants regenerated from the initial cell population of interest.

20

EXAMPLE II

Production of promoters with the isolation method

Construction of adapted genomic libraries from alfalfa
genomic DNA

Adapted genomic libraries from alfalfa DNA were made by using the Universal GenomeWalker™ kit (Clontech Laboratories, cat # K1807-1). Briefly construction of DNA libraries begins with isolation of very clean genomic DNA that has a very high average molecular weight. The starting DNA must be of considerably higher quality than the minimum suitable for Southern blotting or conventional PCR. Five separate aliquots are then

- thoroughly digested with four different restriction enzymes (EcoRV, DraI, PvuII, ScaI, and SspI) that recognize a 6-base site, leaving blunt ends. Following digestion, each pool of DNA fragments is ligated to the
5. GenomeWalker™ adapter. The same adapted libraries can be used to isolate independent promoter fragment using the adapter primer and gene-specific primers.

Materials and Methods

Promoter isolation by genome walking

10 Nitrite reductase (Nir)

- The coding sequence of the nitrite reductase gene (Nir) from alfalfa was obtained (SEQ ID NO:1). Two gene specific primers were designed (GSP1, 5'-TTGTCACATCAGCACATCCGTCTTTGC-3' (SEQ ID NO:7)); GSP2, 5'-
- 15 TCGCCAAGTATCTTGTTTGAGCACTTG-3' (SEQ ID NO:8)) in the direction C-terminal to N-terminal. The GSP1 primer is located downstream of GSP2 in the coding sequence. Genome walking was performed according to the user manual guide and a unique 4 kb DNA fragment was
- 20 obtained from the PvuII-adapted genomic library. This fragment was subcloned into the vector pGEM-t (Promega, cat# A1360). DNA sequencing of this fragment revealed that it contained both the adaptor primer AP2 and the Nir gene specific primer GSP2 sequences (SEQ ID NO:2).
- 25 The DNA sequence found upstream of the GSP2 primer in the Nir coding sequence was also found in the DNA fragment isolated by genome walking confirming that it corresponded to the Nir gene promoter. The isolated Nir gene promoter sequence consisted of 2860 bp
- 30 upstream of the starting ATG.

In addition to the isolation of the 5' non

- 37 -

conding region (promoter), the genome walking protocol was also used to isolate and clone the 3' non coding sequence (terminator) of the Nir gene. Two other Nir gene specific primers were designed (GSP1' 5'-
5 ATGTCTTCCTTCTCAGTACGTTTCCTC-3' (SEQ ID NO:9)); GSP2' 5'-
CAAGTTGATGCATCAAGGTGGATCCTAGA-3' (SEQ ID NO:10)) and used to PCR amplify a Nir specific fragment from the alfalfa adapted genomic libraries. A 3.5 kb DNA
10 fragment was amplified from the EcoRV-adapted library, cloned into vector pGEM-t (Promega), and sequenced (SEQ ID NO:3).

Plastocyanin

Simultaneously to the isolation of nitrite reductase promoter, a second promoter was isolated
15 using the Universal GenomeWalking™ kit. A fragment of the coding sequence of the alfalfa plastocyanin gene was obtained (SEQ ID NO:3). From this sequence, two gene-specific primers (GSP1, 5'-
AGGAGCATTGAGAAGATCTTCTTCAGG-3' (SEQ ID NO:11)); GSP2,
20 5'-GCTGCATCAACCCCGCTTGAATCTCG-3' (SEQ ID NO:12)) were designed. Genome walking was performed according to the user manual guide and a unique 0.7 kb DNA fragment was amplified from the ScaI-adapted genomic library. This fragment was subcloned into the vector pGEM-t
25 (Promega, cat# A1360). DNA sequencing of this fragment revealed that it contained both the adapter primer AP2 and the plastocyanin gene specific primer GSP2 sequences (SEQ ID NO:4). Furthermore, the 3' end putative plastocyanin promoter sequence had complete
30 DNA sequence homology with the 5' end of the plastocyanin coding region used to design the gene specific primers. In addition, the isolated DNA fragment included the predicted starting codon (ATG) of

the plastocyanin gene. This confirmed that the DNA sequence obtained by genome walking was the plastocyanin gene promoter. The isolated plastocyanin promoter was 517 bp long (SEQ ID NO:5).

- 5 The plastocyanin terminator was also identified and cloned using the genome walking protocol. Two plastocyanin gene specific primers were designed (GSP1' 5'-GCGTTACTTTGGATGCTAAGGGAACCT-3' (SEQ ID NO:13)); (GSP2' 5'-TCACGCAGGAGCTGGTATGGTTGGACA-3' (SEQ ID NO:14))
- 10 and used to PCR amplify a plastocyanin specific fragment from the alfalfa adapted genomic libraries. A 1.3 kb DNA fragment was amplified from the *Stu*I-adapted library, cloned into vector pGEM-t (Promega), and sequenced (SEQ ID NO:6).

15 Construction of promoter :reporter gene constructs
 Nitrite reductase (Nir)

- This step was performed using a ligation by amplification protocol developed by Darveau et al. (Methods in Neuroscience, 26 : 77-87). A 2 kb fragment
- 20 of the Nir promoter was fused to the B-glucuronidase (GUS) reporter gene. Four PCR primers were used; a Nir promoter specific primer (5'-GATCTCCCTAACAGTCTCAAAAGTGT-3' (SEQ ID NO:15)), a Nir-GUS ligation specific primer (5'-GGTTTCTACAGGACGTAACATTTTGGAGAAGAGAGTGTGTTTGG-3' (SEQ ID
- 25 NO:16)), a GUS ATG primer (5'-ATGTTACGTCCTGTAGAAACC-3' (SEQ ID NO:17)), and a Nopaline Synthase (NOS) terminator primer (5'-GCCATGAATTCCCGATCTAGTAACATAG-3' (SEQ ID NO:18)). The PCR amplification was performed
- 30 in a single reaction using two template DNA (the pGEM-T plasmid containing the 4 kb Nir promoter insert and the binary vector pBI221 containing the GUS-NOS DNA

- 39 -

fragment). The resulting Nir-GUS PCR fragment was digested SacI-XmaI and subcloned directly into the vector pBI201. This vector (pBI201) is the result of the insertion of the EcoRI-HindIII DNA fragment from the binary vector pBI101, which contain the promoterless GUS reporter gene linked to the NOS terminator, into the vector pBluescript. Finally, the EcoRI-HindIII fragment from the resulting plasmid was isolated and inserted into the EcoRI-HindIII restriction site of the binary vector pBI101. This construct contains the Nir promoter and the NOS terminator.

A construct containing the Nir promoter and the Nir terminator was also made. To allow the cloning of the Nir terminator in fusion to the Nir promoter-GUS fusion, the Nir terminator (SEQ ID NO:3) was PCR amplified using the following primers : Sac primer 5'-AGAAGAGCTCTTGTACATTGATAAGTCA-3' (SEQ ID NO:19), Eco primer 5'-AGAAGAATTCGTTTCCCGATACTTCAACT-3' (SEQ ID NO:20). The resulting PCR fragment was digested SacI-EcoRI and subcloned into the binary vector containing Nir-promoter-GUS-Nos construct in the same restriction sites.

Plastocyanin

For analysis of the expression pattern of the plastocyanin promoter in plant, it was fused to the GUS reporter gene similarly to the Nir promoter. The 517 bp plastocyanin promoter isolated previously was fused to the GUS gene via the ligation by amplification protocol. The resulting PCR fragment containing the plastocyanin promoter fused to the GUS reporter gene was sequenced and subcloned into the vector pBI201

using restriction digest SacI-XmaI. Similarly to the Nir promoter construct, the EcoRI-BamHI DNA insert from the resulting plasmid was reinserted into binary vector pBI101.

- 5 For the plastocyanin promoter GUS plasto-terminator fusion, the plastocyanin terminator was amplified by PCR using two primers containing either a SacI or a EcoRI restriction sites (SacI primer 5'-AGAAGAGCTCGTTAAATGCTTCTTCGTCTCCTA-3' (SEQ ID NO:21));
- 10 EcoRI primer 5'-AGAAGAATTCTCCTTCCTAATTGGTGTACTATCA-3' (SEQ ID NO:22)). The template DNA used for this PCR was the plasmid containing the DNA fragment obtained by genome walking toward the 3' end of the plastocyanin cDNA. The resulting PCR fragment was digested SacI-
- 15 EcoRI and subcloned into the binary vector containing the Plasto-promoter-GUS-NOS construct using the same restriction sites.

Plant transformation

- The recombinant plasmids were introduced into
- 20 *Agrobacterium tumefaciens* strain LBA4404 by electroporation as described in Khoudi et al (1999, Biotechnology and Bioengineering 64:135-143). *Agrobacterium*-mediated plant transformation was performed according to Horsch et al, (1985, Science
- 25 227:1229-1231). Briefly, selected strains were co-cultivated with tobacco leaf disks for 2 days on MS medium without kanamycin. After this period, the explants were transferred to the selection medium (MS with Kanamycin). The explants were kept on this medium
- 30 for 3 weeks to allow the formation of calli and shoots from the transfected cells. The kanamycin resistant shoots were transferred into the rooting MS medium.

Rooted plantlets were transfer to soil and grown to maturity in the greenhouse. Integration of the transgene was verified by PCR amplification NptII gene using specific primers. Several independent transgenic
5 plants from each different constructs were generated.

Results

Promoter activity analysis

Nitrite Reductase

For promoter activity analysis of the T0
10 transgenic plants containing the Nir-GUS constructs, following rooting in-vitro, plants were transferred to vermiculite and allowed to grow for three weeks in the greenhouse. To test the effect of nitrate (a know inducer of the Nir genes) on expression patterns
15 controlled by the Nir promoter, plants were given a particular nitrogen diet. For the first three weeks in vermiculite, plants were watered with Hoagland solution; 2mM KH_2PO_4 , 2mM $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.55 mM K_2SO_4 , 15mM KCl , 10 mM NH_4Cl , 2.8 mM $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.05 mM Nafe EDTA
20 (Hoagland and Arnon, 1950) and 1 ml of micronutrients (1g/L H_3BO_3 , 1g/L $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$, 0.58 g/L $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 0.13 g/L $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0.1 g/L $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$). Then for an extra week, plants were watered with the same media except that the ammonium chloride had been replaced by 40 mM
25 potassium nitrate(KNO_3). At the end of this four week treatment, the third leaf from the top of each transgenic plants was harvested and analyze for b-glucuronidase activity according to Jefferson et al. (1987). This source of plant material (third leaf from
30 top) was chosen as it represent a nearly mature leaf that still retains its full metabolic capacity. As non-induced controls, the third leaf from plants after

the initial three weeks of water treatment were harvested and analyzed for GUS activity. In addition, plants containing a cauliflower mosaic virus (CaMV) 35S-GUS-NOS construct were also used as control.

5 Results from Nir promoter activities are shown in Fig. 3.

While CaMV promoter shows no induction by the KNO_3 treatment with a median GUS activity at around 4 nmol MU/mg protein/min with or without treatment, the

10 NIRpro-NOS construct responded significantly to the same treatment with median GUS activity at 0.5 and 4 nmol MU/mg protein/min without and with nitrate induction respectively. This represents an eight fold induction. In addition, an almost three fold increase

15 in GUS activity is seen when the NOS terminator is replaced by the NIR gene terminator. This increase is seen in both non-induced and induced conditions indicating that the NIR terminator contains important regulatory sequences require to obtain maximum

20 efficiency of gene expression under the control of the endogenous NIR gene regulatory environment.

Plastocyanin

For promoter activity analysis under the control of the plastocyanin gene promoter, T0 plants

25 were transferred to vermiculite following rooting in-vitro and allowed to grow for a three week period in the greenhouse. Since the plastocyanin gene is not induced by nitrate, no particular water treatment was used. The third leaf from the top of three week old

30 greenhouse transgenic plants was harvested and analyze for GUS activity. Two different promoter :terminator constructs were tested (plastopro-Nos ter, plastopro-

plastoter). In addition, plants containing the CaMV35S -GUS constructs were used as control. Fig. 4 shows the differences in GUS activity between these three populations of transgenic plants. There is a 15-
5 20 X increase in gene expression between the plants containing the 35S-GUS and the plastopro-Nos ter constructs. In addition, similarly to the GUS activity under the control of Nirpro-Nirter constructs, GUS activity with the plastopro-plastoter constructs is 2X
10 higher than the GUS activity with the plastopro-nos ter construct. Again this indicates, that the plasto terminator must contain important regulatory sequences require to maximize gene expression under the control of the endogenous plasto gene regulatory environment.
15 It is also important to note that the proposed system enable the evaluation of promoter activities of a broad range of expression level from a low expression (<1 nmol MU/mg protein/min for uninduced NIRpro-Nos construct) to a very high expression (180 nmol MU/mg
20 protein/min for the plastopro-plastoter construct) (Fig. 5).

Production of small quatity of proteins in cell cultures

One of the direct applications of these
25 expression regulatory sequences is to regulate expression of genes of interest for molecular farming purpose. Therefore, the plastocyanin promoter and the Nitrite reductase promoter were fused to a gene of interest coding for a polypeptide of 34 kd to give five
30 separate constructs (numbered 8, 11, 19, 23, and 24). Constructs 8, 11 and 19 contain the gene of interest fused to the nitrite reductase promoter while constructs 23 and 24 use the plastocyanin promoter to

drive the gene of interest. These constructs were inserted by triparental mating in *Agrobacterium tumefaciens* LB4404 and used to inoculate alfalfa petioles for *Agrobacterium* mediated transformation protocol adapted specifically for alfalfa according to Daniel Brown, Research Scientist Agriculture and Agro-Food Canada, London Station, Ontario (personal communication). Following *Agrobacterium* transformation, plant petioles were cocultured in B5H solid media (Tian et al., 2000, Can. J. Plant Sci 80:765-771) without selectable marker for a period of 2 days under low light conditions (16-h photoperiod with photosynthetic Photon Flux of about 50 $\mu\text{mol m}^{-2} \text{s}^{-1}$ at 25°C). Then plant material was transferred to B5H solid media with 75 mg/L kanamycin for a period of 4-6 weeks to allow callus formation under low light conditions.

For the expression in cell culture, Calli growing on selective media were desegregated manually and transferred in liquid modified B5 media (Tian et al., 2000, Can. J. Plant Sci 80:765-771) with 24 mg/L kanamycin at a ratio of 0.5-1g plant material/20 ml culture media. Liquid cultures were shaken at 120 RPM under low light conditions for 2-3 weeks. Following this, crude protein extracts were obtained from the cell suspensions. Cell cultures were centrifuged and the pellets were ground in mortar and pestle using liquid nitrogen and sand. The resulting pulverized plant material was mixed with extraction buffer (50 mM NaHPO₄, pH 7.0, 10 mM EDTA, 0.1% Triton X-100) and centrifuged at 21000Xg for 10 minutes to pellet cell debris. The supernatant was collected and precipitated with trichloroacetic acid (TCA). This precipitation

was performed by adding 1 volume of TCA 10%, mixing by vortex and allowing the protein to precipitate for 30 minutes on ice. The mixture was centrifuged at 10000Xg for 15 minutes. The supernatant was removed by
5 decanting immediately and removing the remaining liquid by aspiration. The pellet was resuspended in 10 μ l of sample lysis buffer 1X and 5 μ l of 1.5M Tris, pH 8.8 was added to neutralize the sample before electrophoresis.

10 For Western blot analysis, samples were separated by SDS-PAGE on 12% acrylamide and electrotransferred onto polyvinylidene difluoride (PVDF) membrane. Membrane blocking and detection of conjugated-peroxidase activity were performed with the
15 chemiluminescence kit (Boehringer Mannheim), as described by the manufacturer. Primary antibody and horseradish peroxidase labeled secondary antibody were diluted at the optimal dilution in 0.5% blocking buffer. Figure 5 shows detection of protein X in
20 alfalfa cell cultures by Western.

This result demonstrates that alfalfa cell cultures can express the gene of interest and produce its corresponding polypeptide of 34 kd in quantities enough for detection by western blot analysis. The
25 single band observed when the protein of interest is expressed in alfalfa cell culture indicates that a single form of the 34 kd protein is found in transgenic plant cells.

While the invention has been described in connection with specific embodiments thereof, it will be
30 understood that it is capable of further modifications and this application is intended to cover any varia-

tions, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the present disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features herein before set forth, and as follows in the scope of the appended claims.

WHAT IS CLAIMED IS:

1. A method for isolating and characterizing an expression regulatory sequence for expression of recombinant polypeptide and/or RNA comprising the steps of:

- a) isolating mRNA from a cell;
- b) preparing a cDNA library from said mRNA;
- c) producing at least one oligonucleotide primer from cDNAs of said cDNA library of step b), said oligonucleotide primer allowing amplification of genomic sequences upstream or downstream of said cDNAs;
- d) performing amplification of said genomic sequences upstream or downstream of said cDNAs with said oligonucleotide primer of step c) on a genomic sample;
- e) linking said amplified sequence of step d) to a gene encoding for a detectable polypeptide and/or RNA to form a DNA expression vector for expression of said detectable polypeptide and/or RNA; and
- f) selecting an expression regulatory sequence of a vector of step e) by measuring the level of expression of said detectable polypeptide and/or RNA under conditions allowing activation of said expression regulatory sequence and expression of said detectable polypeptide and/or RNA.

- 48 -

2. The method of claim 1, wherein said cell of step a) is a plant cell.
3. The method of claim 1, wherein said cell of step a) is an alfalfa cell.
4. The method of claim 1, wherein origin of said gene encoding polypeptide and/or RNA origins from the group consisting of a animal, a mammal, a plant, an insect, a yeast, a mold, a bacterium, and a virus.
5. The method of claim 1, wherein said polypeptide and/or RNA is selected from the group consisting of a pharmaceutical, an agronomical, an environmental, an industrial, a nutraceutical, a cosmeceutical a polypeptide, a gene product marker, a fusion protein, green fluorescent protein, and a β -glucuronidase.
6. The method of claim 1, wherein said condition of step f) is an *in vitro* or an *in vivo* condition.
7. The method of claim 6, wherein said *in vitro* conditions allows for expression of detectable polypeptide and/or RNA from a transitory transfected cell, a stably genetically transformed cell, or in a reaction buffer.

8. The method of claim 6, wherein said *in vivo* expression is expression in a cultured cell, or a growing organism.

9. The method of claim 1, wherein said polypeptide and/or RNA is indirectly detected by using at least one of antibodies, Western blot, Northern blot, *In situ* hybridization, colorimetry, optical densitometry, spectrophotometry, or electrophoresis.

10. The method of claim 1, wherein said polypeptide and/or RNA comprises a tag to be directly detected or for purification of said polypeptide.

11. The method of claim 10, wherein said tag is a self-cleavable tag.

12. The method of claim 1, wherein said genomic sequences comprise expression regulatory sequence which are further sequenced.

13. The method according to claim 1, wherein said genomic sequences is an expression regulatory sequence which is natively located upstream or downstream of a gene encoding a polypeptide and/or RNA and controls the expression of said gene encoding a polypeptide and/or RNA.

14. A method of producing adapted DNA expression vector for expression of recombinant polypeptides and/or RNA comprising the steps of:

- a) isolating mRNA from a cell;
- b) preparing a cDNA library from said mRNA;
- c) producing at least one oligonucleotide primer from cDNAs of said cDNA library of step b), said oligonucleotide primer allowing amplification of genomic sequences upstream or downstream of said cDNAs;
- d) performing amplification of said genomic sequences upstream or downstream of said cDNAs with said oligonucleotide primer of step c) on a genomic sample;
- e) linking said amplified sequence of step d) to a gene encoding for a detectable polypeptide and/or RNA to form a DNA expression vector for expression of said detectable polypeptide; and
- f) selecting a DNA expression vector of step e) by measuring levels of expression of said detectable polypeptide and/or RNA under a condition allowing activation the expression of said detectable polypeptide and/or RNA.

15. The method of claim 14, wherein said cell of step a) is a plant cell.

16. The method of claim 14, wherein said cell of step a) is an alfalfa cell.

17. The method of claim 14, wherein origin of said gene encoding polypeptide and/or RNA origins from the group consisting of an animal, a mammal, a plant, an insect, a yeast, a mold, a bacterium, and a virus.

18. The method of claim 14, wherein said polypeptide is selected from the group consisting of a pharmaceutical, an agronomic, an environmental, an industrial, a nutraceutical, and a cosmeceutical polypeptide, or a gene product marker, a fusion protein, a green fluorescent protein, and a β -glucuronidase.

19. The method of claim 14, wherein said condition of step f) is an *in vitro* or an *in vivo* conditions.

20. The method of claim 19, wherein said *in vitro* condition allows for the expression of a detectable polypeptide in a transitory transfected cell, a stably genetically transformed cell, or in a reaction buffer.

21. The method of claim 19, wherein said *in vivo* expression is expression in a cultured cell, or in a growing organisms.

22. The method of claim 14, wherein said polypeptide is indirectly detected by using at least one of antibodies, Western blot, Northern blot, *In situ* hybridization, colorimetry, optical densitometry, spectrophotometry, or electrophoresis.

23. The method of claim 14, wherein said polypeptide comprises a tag to be directly detected or for purification of at least one of said polypeptide or RNA.

24. The method of claim 23, wherein said tag is a self-cleavable tag.

25. The method of claim 14, wherein said DNA expression vector is further sequenced.

26. The method according to claim 14, wherein said genomic sequence comprises at least one expression regulatory sequence which is natively located upstream or downstream of a gene encoding a polypeptide and/or RNA and which control the expression of said gene encoding a polypeptide and/or RNA.

27. A transgenic plant regenerated from stably genetically transformed cell of claim 7 or 20.

28. The method of claim 14, wherein said DNA expression vector comprising said a genomic sequence which comprises an expression regulatory sequence.

29. A DNA expression vector of claim 28 which is a plasmid vector.

30. A DNA expression vector of claim 28 which is a viral vector.

31. A plant cell transformed with the DNA expression vector of claim 29 or 30.

32. A transgenic plant regenerated from the plant cell of claim 31.

33. A method of producing recombinant polypeptides and/or RNA using a plant cell of claim 7, 20, 27, or 31 and/or said transgenic plant of claim 32.

34. A method of isolating and characterizing an expression regulatory sequence for expression of a recombinant polypeptide and/or RNA comprising the steps of:

- a) producing at least one oligonucleotide primer from a cDNA, genomic DNA fragment or synthetic DNA sequence, said oligonucleotide primer allowing amplification of a genomic sequence upstream or downstream of a genomic complementary site of said oligonucleotide primer in a genomic DNA sample;
- b) performing amplification of said genomic sequence upstream or downstream of said genomic complementary site of said oligonucleotide primer a) on a genomic DNA sample;

c) linking an amplified sequence obtained from the amplification of step b) to a gene encoding for a detectable polypeptide and/or RNA to form a DNA expression vector for expression of said detectable polypeptide and/or RNA; and

d) selecting at least one expression regulatory sequence from said vector of step c) by measuring levels of expression of said detectable polypeptide and/or RNA under a condition allowing activation of said expression regulatory sequence and expression of said detectable polypeptide and/or RNA.

35. The method of claim 34, wherein said cell of step a) is a plant cell.

36. The method of claim 34, wherein said cell of step a) is an alfalfa cell.

37. The method of claim 34, wherein origin of said gene encoding polypeptide and/or RNA origins from the group consisting of a animal, a mammal, a plant, an insect, a yeast, a mold, a bacterium, and a virus.

38. The method of claim 34, wherein said polypeptide and/or RNA is selected from the group consisting of a pharmaceutical, an agronomical, an environmental, an industrial, a nutraceutical, a cosmeceutical a

polypeptide, a gene product marker, a fusion protein, green fluorescent protein, and a β -glucuronidase.

39. The method of claim 34, wherein said condition of step f) is an *in vitro* or an *in vivo* condition.

40. The method of claim 39, wherein said *in vitro* conditions allows for expression of detectable polypeptide and/or RNA from a transitory transfected cell, a stably genetically transformed cell, or in a reaction buffer.

41. The method of claim 39, wherein said *in vivo* expression is expression in a cultured cell, or a growing organism.

42. The method of claim 34, wherein said polypeptide and/or RNA is indirectly detected by using at least one of antibodies, Western blot, Northern blot, *In situ* hybridization, colorimetry, optical densitometry, spectrophotometry, or electrophoresis.

43. The method of claim 34, wherein said polypeptide and/or RNA comprises a tag to be directly detected or for purification of said polypeptide.

44. The method of claim 43, wherein said tag is a self-cleavable tag.

45. The method of claim 34, wherein said genomic sequences comprise expression regulatory sequence which are further sequenced.

46. The method according to claim 34, wherein said genomic sequences is an expression regulatory sequence which is natively located upstream or downstream of a gene encoding a polypeptide and/or RNA and controls the expression of said gene encoding a polypeptide and/or RNA.

47. A method of producing an adapted DNA vector for expression of recombinant polypeptides and/or RNA comprising the steps of:

- a) producing at least one oligonucleotide primer from a cDNA, a genomic DNA fragment or a synthetic DNA sequence, said oligonucleotide primer allowing amplification of a genomic sequence upstream or downstream of a genomic complementary site of said oligonucleotide primer in a genomic DNA sample;
- b) performing the amplification of at least one of said genomic sequence upstream or downstream of said genomic complementary site with said oligonucleotide primer of step a) on a genomic DNA sample;
- c) linking an amplified sequence obtained from the amplification of step b) to a gene encoding for a detectable polypeptide and/or RNA to form a DNA expression vector for

expression of said detectable polypeptide and/or RNA; and

- d) selecting a DNA expression vector of step c) by measuring the level of expression of said detectable polypeptide and/or RNA.

48. The method of claim 37, wherein said cell of step a) is a plant cell.

49. The method of claim 37, wherein said cell of step a) is an alfalfa cell.

50. The method of claim 37, wherein origin of said gene encoding polypeptide and/or RNA origins from the group consisting of an animal, a mammal, a plant, an insect, a yeast, a mold, a bacterium, and a virus.

51. The method of claim 37, wherein said polypeptide is selected from the group consisting of a pharmaceutical, an agronomic, an environmental, an industrial, a nutraceutical, and a cosmeceutical polypeptide, or a gene product marker, a fusion protein, a green fluorescent protein, and a β -glucuronidase.

52. The method of claim 37, wherein said condition of step f) is an *in vitro* or an *in vivo* conditions.

- 58 -

53. The method of claim 52, wherein said *in vitro* condition allows for the expression of a detectable polypeptide in a transitory transfected cell, a stably genetically transformed cell, or in a reaction buffer.

54. The method of claim 52, wherein said *in vivo* expression is expression in a cultured cell, or in a growing organisms.

55. The method of claim 37, wherein said polypeptide is indirectly detected by using at least one of antibodies, Western blot, Northern blot, *In situ* hybridization, colorimetry, optical densitometry, spectrophotometry, or electrophoresis.

56. The method of claim 37, wherein said polypeptide comprises a tag to be directly detected or for purification of at least one of said polypeptide or RNA.

57. The method of claim 56, wherein said tag is a self-cleavable tag.

58. The method of claim 37, wherein said DNA expression vector is further sequenced.

59. The method according to claim 37, wherein said genomic sequence comprises at least one expression regulatory sequence which is natively located upstream or downstream of a gene encoding a polypeptide and/or

RNA and which control the expression of said gene encoding a polypeptide and/or RNA.

1/5

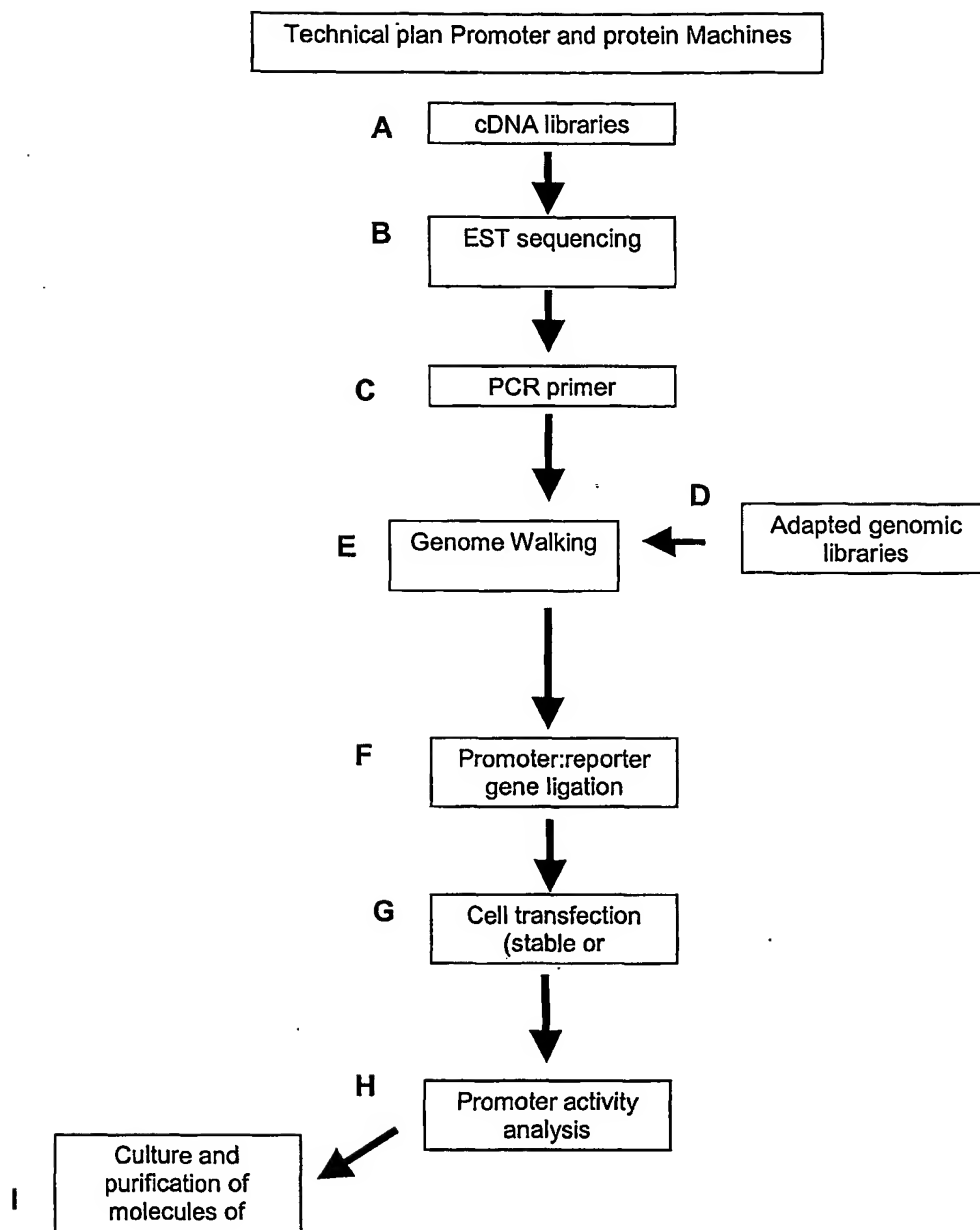


Fig. 1

2/5

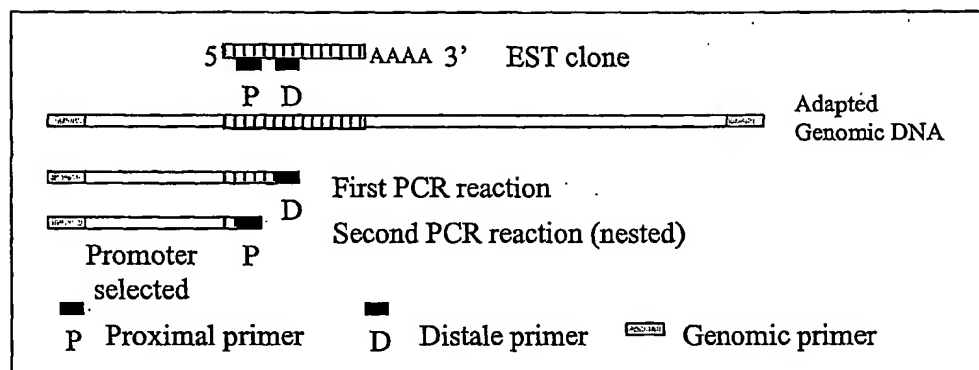


Fig. 2

3/5

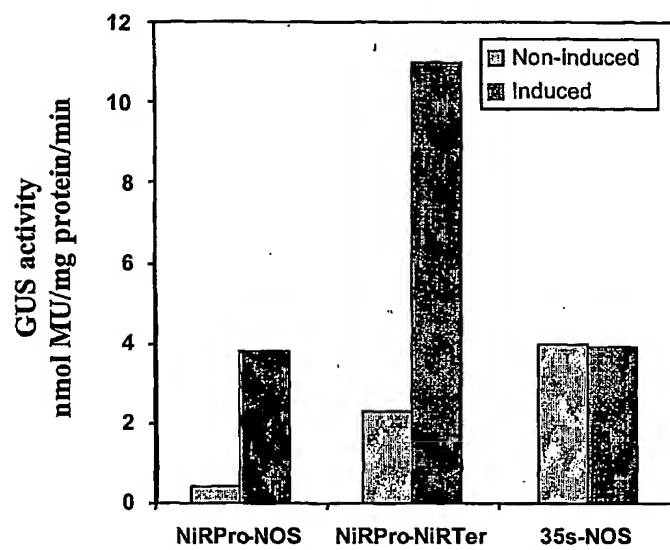


Fig. 3

4/5

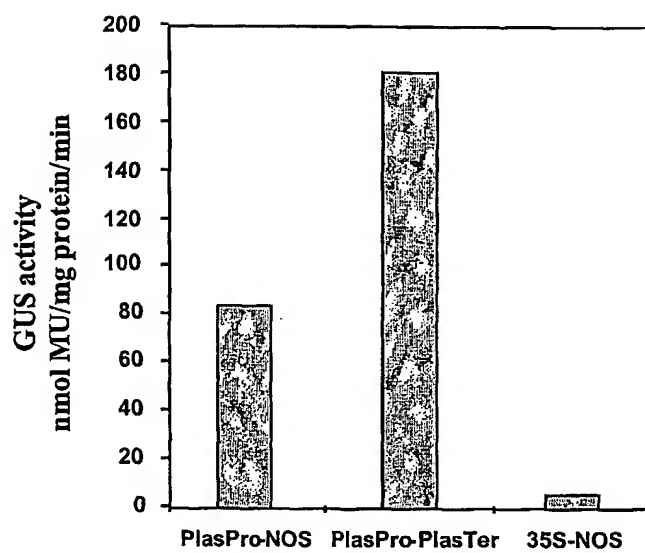


Fig. 4

5/5

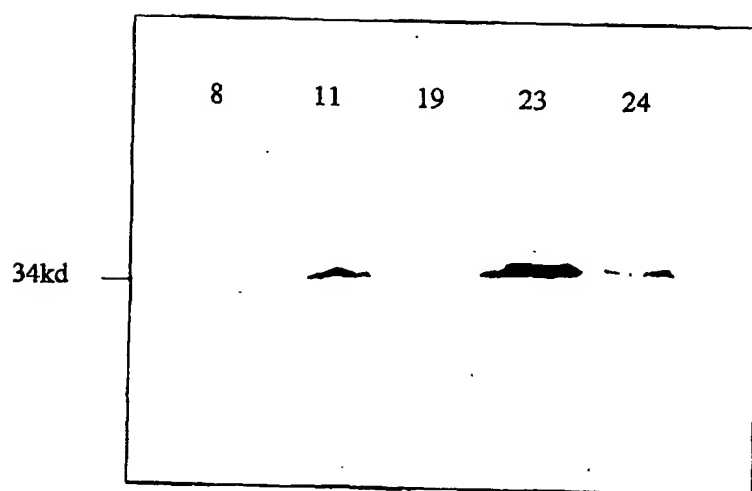


Fig. 5

SEQUENCE LISTING

<110> VÉZINA, Louis-Philippe
D'AOUST, Marc-André
ARCAND, François
BILODEAU, Pierre
MEDICAGO inc.

<120> METHOD OF SELECTING PLANT PROMOTERS TO
CONTROL TRANSGENE EXPRESSION

<130> 14149-7PCT

<140> 60/244,214

<141> 2000-10-31

<160> 22

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1752

<212> DNA

<213> Artificial Sequence

<220>

<223> NIR promoter

<400> 1

atgtcttcct	tctcagtagc	tttcctcact	ccaccatcca	tctctcgtcc	caacaaaaca	60
tggctactat	ctgctgcaac	tccatcagtt	gcacctgttt	caacaccaca	agttgatgca	120
tcaagggttg	atcctagagt	tgaggaaaaa	gatggttact	gggttttgaa	ggaagagtat	180
agaggaggta	ttaatcctca	ggagaaagtt	aagattcaga	aagaacctat	gaagcttttt	240
atggaagggtg	ggattaatga	tttggctaata	atgtctcttg	aagagattga	aagctctaag	300
cttactaaag	atgatattga	tgtagactt	aaatggcttg	gtctttttca	tagaaggaaa	360
catcattatg	gtagatttat	gatgagactg	aaacttccaa	atggggtaac	aacaagtgtc	420
caaacaagat	acttggcgag	tgtgataaaa	aaatatggca	aagacggatg	tgctgatgtg	480
acaacgaggc	agaattggca	aattcgaggt	gtaacgttac	ctgatgtccc	tgaaattcct	540
aagggccttg	cagaggctcg	cttgacaagt	ctgcagagtg	gaatggacaa	tggtcgaaac	600
ccagttggta	accctcttgc	tggtattgac	cctgatgaga	ttgttgatac	aagaccttac	660
accaatttgc	tgtcccaatt	catcactgct	aattcacttg	gtaatccaac	cattacaaac	720
ttgccaagga	agtggaaatgt	atgtgtgata	ggttcccatg	atcttttcga	gcacccgat	780
attaacgata	ttgcttata	gcctgctaata	aaggatggc	gatttggatt	caacttattg	840
gtgggtgggt	tctttagtcc	caagcgatgt	gctgaagcag	ttccacttga	tgcatgggtc	900
tctgcagatg	atgttatccc	actttgtaaa	gctgtccttg	agacctatag	ggacctcggc	960
acaagaggga	atagacagaa	aaccagaatg	atgttgggtga	tcgatgaact	tggtatagaa	1020
gtattcagat	cagagggtgga	aaaaagaatg	ccagagaaga	agctagagag	agcatccaaa	1080
gaagaacttg	tccaaaaaca	atgggaaaga	ggagacatct	taggtgttca	tccacaaaaa	1140
caagaagggt	taagctatgt	tgggaattcac	attccagttg	gtagaatcca	agtagatgag	1200
atggatgagc	tagctcgtat	tgccgacgaa	tacggaaaccg	gcgaactaag	gctaaccgta	1260
gagcaaaaaca	taataattcc	aaatgtggaa	aactcaaaac	ttgatgcatt	gctaaatgaa	1320
cctctcttga	aagacaaatt	ctcaccagaa	ccttccatcc	taatgaaaac	acttgtggca	1380
tgcaactggt	accaattttg	tggccaagca	ataattgaaa	caaaacaaag	agctttaaaa	1440
gtaactgaag	aagttgagag	atatgtggct	gtgagcaaac	cagtgagaat	gcattggact	1500
ggttgtccta	acacatgtgg	tcaagttcag	gttgctgata	ttggttttat	gggttgatat	1560

gctagggatg	aaaatggtaa	ggctactgaa	ggtgttgata	ttttccttgg	tgggagaatt	1620
ggaagtgatt	ctcatttagc	tgaggtgtat	aagaaagggtg	tcccttgcaa	ggacttgggtg	1680
cctattgtag	ctgacatttt	ggttaaataat	tttgagactg	tccaaaggaa	tagagaagaa	1740
ggggatgatt	aa					1752

<210> 2

<211> 3775

<212> DNA

<213> Artificial Sequence

<220>

<223> AP2 adaptor with NIR promoter

<400> 2

gggaattcga	ttactatagg	gcacgcgtgg	tcgacggccc	gggctgggtct	gtacattcat	60
cttgccgcct	ttgcattcac	ttggccacaa	agagtagaga	gaaggaagag	aagagcccag	120
acttcaagaa	gcgaccttgc	aagtgcactc	gagggtcaga	aactgtatat	catatctatg	180
tgagagaaag	gggaacattt	gagatggagt	ccatttactt	gaggtatact	tattattttg	240
atcaataaat	ttgtatactt	cttatttaga	tcaataaatt	tgtcattaag	ctataatcca	300
aaataaatta	cgatcaaata	tgcaaatgtt	agccagtact	tgtgttaaac	ttgatggcat	360
ctcttggttt	ctttggcaat	cacatgccta	agaaataaat	agtatcatat	gatttgtgtt	420
ggtcagactt	cagagtcaga	tgactctgtt	tgataaaaca	gcttaattaa	gcgcttatag	480
aatatcatat	gatttgtgtt	ggtcagactt	cagagcatct	cttggtttct	ctggcaatca	540
tatgcctaag	aaataaatag	tatcatatga	ttgtgttttg	tcagacttca	gagtcagatg	600
accctgtttg	ggtaaacagc	cttaattaagt	gcttatagaa	taagcgctta	tcataataag	660
gcttttgtac	agttatttct	atgaaagtag	aagaaatagt	catattgttt	taatataagc	720
tatcctggag	agcttgtgga	aataaccaga	aaagaactta	tggacacgtc	atgagctggt	780
tacataagat	ctccctaaca	gtctcaaaag	tgtttatgcc	agtagataaa	ttcaaataag	840
tcaatctaaa	cagaccctaa	atccattatg	gtacctatca	ttttagctta	ttccatcttt	900
attaagaatg	tcagagata	acataatgat	aacacattat	tttgacacaa	atgggcagat	960
ctagcaattt	aactctggag	tccttcaaga	ctgctgttct	tacgaagttc	acgtccctga	1020
atcatgttcc	tgtatggaag	cctgaaagac	ctcaaattct	aaaagggtggc	gataaattga	1080
aggtttacia	aatataccct	gcgggcttga	cacagaggca	agctctttat	accttccagt	1140
tcaacgggga	tgttgatttc	agaagtcact	tggagagcaa	tccttgtgcc	aagtttgaag	1200
taatttttgt	gtagcatatg	ttgagctacc	tacaatttac	atgatcacct	agcattagct	1260
ctttcactta	actgagagaa	tgaagtttta	ggaatgagta	tgaccatgga	gtcggcatgg	1320
ctttgtaatg	cctaccctac	tttggccaac	tcacggggga	tttacattca	gaaaatatac	1380
atgacttcaa	ccatacttaa	accccttttt	gtaagataac	tgaatgttca	tatttaattg	1440
tgggtgtgac	tgatttatat	tgatttatat	cagacagtta	caagttggac	aacaagattg	1500
tgggtctgta	ctgttattta	tttatttttt	tttttagaga	aacaccttat	cttttgtttc	1560
gtttgaatgt	agaatgaaaa	taaaagaaag	aaaatataac	atcatcggcc	gcgcttgtct	1620
aatttcgggc	agtttaggatc	ctctccggtc	accggaaagt	ttcagtagaa	gaaacaaaac	1680
accgtgacta	aatgataact	attatttttt	ttatttgtgt	tttctttttt	ctaccggaac	1740
tttttagaac	ggatcccaac	tcgttccggg	gccgctacaa	ctgaaacaaa	agaagatatt	1800
ttctctctct	tcagaaatgt	aagttttcct	ttacagatac	ccattcacca	tttgattcag	1860
atgtgtgtgac	tagagataaa	gcataactaa	ttgactcttg	gaaaccata	aagtttatgt	1920
tatccgtgtt	ctggaccaat	ccacttgggg	gcataacctg	tgtctatgtg	tggtttgggt	1980
tccattctga	tttatgcggc	gacttgtaat	ttaaaatcta	ggaggggcag	acattgaaca	2040
atcccaatat	tttaataact	tatgcaagat	tttttttatt	aatgagatga	tgtgtttgtg	2100
actgagattg	agtcatacat	ttcataaga	aatgggtcca	agtaccaaac	tatcatgacc	2160
cagttgcaaa	catgacgttc	gggagtggtc	actttgatag	ttcaatttca	tcttggcttc	2220
ttattccttt	tataattcta	attcttcttg	tgtaaactat	ttcatgtatt	atttttcttt	2280
aaaatttaca	tgtcatttat	ttgcctcac	attcaactt	ttgcataata	caatgataag	2340
tgatattttg	actcacaaaa	tttatacaca	atttcgacat	cgtttattat	gttcattgga	2400
tgattaacaa	atataacaaa	ctttgcaact	aattaaccac	caactgaata	taattaacta	2460
taactgtgaa	agtagttaac	catatttttt	agatgtatat	atcatccgtt	gaatgtaatt	2520
attcatatat	ttgaactaag	ttaccctaca	acttaaagaa	cttaaagaac	tcggtttgag	2580

acctggggac	gaaaatgtaa	tgagacttta	atgttgactt	tgacaccgca	ccacatgtgc	2640
cttttacata	tagtttatat	gacaagtaat	gacaatcctt	gctctattat	aaggcgaccc	2700
ttagctccaa	ccaaaggacg	atggagttaa	gaaagaaact	cttgcttact	tgtaagggtcc	2760
acacttcctt	actcacctct	caatttcctc	ctacaaaaat	gtccaaactt	ctctttctca	2820
caatcacaaa	ctcattccaa	acacactctc	ttctccaaaa	atgtcttcct	tctcagtacg	2880
tttctccacc	ccaccatcca	tctctcgtcc	caacaaaaa	tggtctactat	ctgctgcaac	2940
tccatcagtt	gcacctgttt	caacaccaca	agttgatgca	tcaagggttg	agcctagagt	3000
tgaggaaaaa	gatggttact	gggttttgaa	ggaagagtat	agagggggta	ttaatcctca	3060
ggagaaagtt	aagattcaga	aagaacctat	gaagcttttt	atggaagggtg	ggattaatga	3120
tttggcta	atgtctcttg	aagagattga	aagctctaag	cttactaaag	atgatattga	3180
tgtagactt	aatggccttg	gtctttttca	tagaaggaaa	catcattgta	agttttttta	3240
ccttcttttt	atacctcaaa	gttctctcat	actctgtatt	tgtttattag	tttttgtaga	3300
cttaaatatt	ctctttgatt	tacatagtga	aactccattt	ttgtttccga	aattgtagtg	3360
tgtatagtct	agaaaattaa	gaagtagaca	aaatgattta	tgagattgta	aattgtaggc	3420
tttttatcaa	tttattaatt	ttagagacca	aaatttgcct	atcttatttg	gaccaatatt	3480
gtatgtcagg	atcgacatga	gtttagtaaa	atcatgacgg	caccatgact	gtgtggaagc	3540
ttctttgtgt	aactttaacc	aaaattatat	ggcacaccat	aattatgcaa	actcacgctc	3600
gatccaaaca	tagaaattcg	gtgttaatct	ttgtgagaat	aaaaagctat	gagttatggt	3660
gtactaattt	atttcattg	tgaaaatcag	atggtagatt	tatgatgaga	ctgaaacttc	3720
caaatggggt	aacaacaagt	gtcaaacaa	gatacttggc	gaaatcacta	gtgaa	3775

<210> 3

<211> 3548

<212> DNA

<213> Artificial Sequence

<220>

<223> Plastocyanin promoter promoter

<400> 3

gcgccgcg	gaattcgatt	caagttgatg	catcaagggt	ggagcctaga	gttgaggaaa	60
aagatggta	ctgggttttg	aaggaagagt	atagaggagg	tattaatcct	caggagaaag	120
ttaagattca	gaaagaacct	atgaagcttt	ttatgggaagg	tgggattaat	gatttggtca	180
atatgtctct	tgaagagatt	gaaagctcta	agcttactaa	agatgatatt	gatgttagac	240
ttaaatggct	tggctctttt	catagaagga	aacatcattg	taagtttttt	tactttcttt	300
ttatacttca	aagttctctc	atactctgta	tttgtttatt	agtttttgta	gacttaaata	360
ttctctttga	tttacatagt	gaaactccat	tcttgtttcc	gaaattgtag	tgtgtatagt	420
ctagaaaatt	aagaagtaga	caaaatgatt	tatgagattg	taaattgtag	gctttttatc	480
aatttattaa	ttttagagac	caaaatttgc	ctactcttatt	tggaccaatt	tattgtatgt	540
taggactgac	atgagtttag	caaaatcatg	acggcaccat	gactgtgttg	aagcttcttt	600
gtgtaacttt	aacccaaatt	atatggcaca	ccatgattat	gcaaactcac	cgtcaatcca	660
aacatagaaa	ttcagtggtta	atctttgtga	caataaaaaa	ctatgagtta	tgttgtagta	720
atttatttcc	attgtgaaac	tcagatggta	gatttatgat	gagactaaaa	ctcccaaattg	780
gggtaacaac	aagtgtctca	acaagatact	tggcgagtgt	gataaaaaaa	tatggcaaaag	840
acggatgtgc	tgatgtgaca	acgaggcaga	attggcaaatt	tcgaggtgta	acgttacctg	900
atgtccctga	aattcttaag	ggccttgca	aggtcggctt	gacaagtctg	cagagtggaa	960
tggacaatgt	tcgaaaccga	gttggttaacc	ctcttgctgg	tattgaccct	gatgagattg	1020
ttgatacaag	accttacacc	aatttgctgt	cccaattcat	cactgcta	tcacttggtta	1080
atccaaccat	tacaaacttg	taagtctaaa	ctatctcabc	tttatatttc	actcattata	1140
tcattattag	agtttagttac	ttgcattgca	agcattacgt	gaccgtgtgt	agcctctaaa	1200
tccttttgat	aatatgtgca	ggccaaggaa	gtggaatgta	tgtgtgatag	gttcccatga	1260
tcttttcgag	catccgcata	ttaacgatct	tgcttatatg	cctgctaata	aggatggctg	1320
atttggattc	aacttatttg	tgggtgggtt	ctttagtccc	aagcgatgtg	ctgaagcagt	1380
tcacttgat	gcatgggtct	ctgcagatga	tgtttatcca	ctttgtaaa	ctgtccttga	1440
gacctatagg	gacctcgcca	caagagggaa	tagacagaaa	accagaatga	tgtgggtgat	1500
cgatgaactt	gtaagttacc	actttttttc	ttcacatatt	attaactgaa	gtgactttaa	1560
cgaccatttt	acaattgaaa	tttaagtggg	ttttagccct	atcattacaa	gaacaaattt	1620

gttaattcac	tagcaagagc	aattccactt	tggcttggac	atgacaagtg	tttgtgaaat	1680
gcaggggata	gaagtattca	gatcagaggt	ggaaaaaaga	atgccagaga	agaagctaga	1740
gagagcatcc	aaagaagaac	ttgtccaaaa	acaatggaaa	gaggagacat	cttaggtggt	1800
catccacaaa	aacaagaagg	tttaagctat	gttgggaattc	acattccagt	tggtagaatc	1860
caagcagatg	agatggaaga	gctagctcgt	atcgccgatg	aatacggaac	cggagaacta	1920
aggctaaccg	tggagcaaaa	cataataatt	ccaaatgtgg	aaaactcaaa	acttgatgca	1980
ttgctaaatg	aacctctctt	gaaagacaaa	ttctcaccag	aaccttccat	cctaataaaa	2040
acacttgtgg	catgcactgg	taaccaattt	tgtggccaag	caataattga	aacaaaacaa	2100
agagctttta	aagtaactga	agaagttgaa	agacatgtgg	ctgtgagcaa	accagtgaga	2160
atgcattgga	ctggttgtcc	taacacttgt	ggtcaagttc	aggttgctga	tattggtttt	2220
atgggttgta	tgggaaggga	tgagaatggt	aaggctactg	aaggtgttga	tattttcctt	2280
ggtgggagaa	ttggaagtga	ttctcattta	gctgaggtgt	ataagaaagg	tgtcccttgc	2340
aaggacttgg	tgcctattgt	agctgatatt	ttggttaaat	attttggagc	tgtccaaagg	2400
aatagagaag	aaggggatga	ttaaagtata	taggtatttg	gtgattttta	ttgcctctac	2460
acaaaattat	tatgttctgt	ccaaaatata	aagtcacaag	ggataattga	gattgagatg	2520
cagcacgcca	cacatgaact	tgtacatttg	gataagtcac	ttttcattgc	tattttataa	2580
gttacacttt	gaattttata	ataaatttta	ttttatttca	aggaccagat	tttataagga	2640
aaccgcta	ctaactatct	ttactcgtaa	tttgtcattt	gagagctacg	gagatcgttg	2700
agtttacgta	tgagtgttta	gtctcacatt	aattatgaat	ggtcaaaatg	ttaaatttat	2760
aagagatgta	atctatatac	ctaattgcatt	aaaaatttgg	atggagatgc	gacgcccccc	2820
ttttttgtgg	tcctgaagta	tagacttggt	gtcgttctg	gtgcactctc	atacttccca	2880
acaaggagaa	aaaactacca	taacaattaa	caaactaaca	tttgttattt	aaaaaaacat	2940
acggatactg	ttttttcccc	atttatttag	aagatgatgg	cttggaattc	aatggctgag	3000
tttatttttt	ttttggtcgg	gagttgaagt	atcgggaaaa	ctaaatatgc	tatgacttta	3060
aacatttgt	tgatataga	ttagttttca	acttacttaa	aaagtggcaa	actagtttag	3120
tggttctctc	ccttccttgt	agttcaagga	acatgggttt	gaactctgtc	caaatttttg	3180
tactttcaat	tatccatact	ttaaaagcta	tataccacat	cattatattc	aagtcaatga	3240
tcatgcgggc	tgccacatta	gcacgatgt	acacattaat	tttaagtggc	atgaacacat	3300
taacatttca	taaaagctat	gtgccagatc	atcattcaag	tctatgcaca	catggtcaac	3360
acattagtac	catttttttt	tttattgttg	atcagattgg	atgtcggtat	tggttgatg	3420
ctacaaactc	aaacaatctc	cagctgttag	agaacgtcga	aatgaatgc	atcacgggtg	3480
cacacttaga	taccagcccg	ggcgcgtcac	cacgcgtgcc	ctatagtaat	cactagtga	3540
ttcgcggc						3548

<210> 4

<211> 504

<212> DNA

<213> Artificial Sequence

<220>

<223> Putative Plastocyanin promoter promoter

<400> 4

atggccaccg	ttacttccac	caccgttgct	attccatcat	tcacaggcct	taaggcaaac	60
gcaagcaaa	gtaatgccat	agctaagggt	ccaacttcaa	cttctcaatt	gccaaggcct	120
tgtgtcagag	cttccctcaa	agactttgga	gttgctgttg	ttgccactgc	tgcaagtgca	180
ttgttagcta	gcaatgccct	tgcagttgaa	gtgttgcttg	gtgctagtga	tgggggtttg	240
gcttttggtc	caaacaattt	cacagtgaac	gctggagaca	ccattacatt	caagaacaat	300
gctggttttc	ctcacaacgt	tatcttcgat	gaagacgaga	ttccaagcgg	ggttgatgca	360
gccaaaattt	ccatgcctga	agaagatctt	ctcaatgctc	ctggggagac	ttacagcggt	420
actttggatg	ctaagggaac	ctacaaattc	tactgttcac	ctcacgcagg	agctggtatg	480
gttggacaag	tcactgttaa	ttaa				504

<210> 5

<211> 940

<212> DNA

<213> Artificial Sequence

<220>

<223> Plastocyanin promoter

<400> 5

ttcactagtg	attactatag	ggcacgcgtg	gtcgcacggcc	cgggctggta	aaattaaaag	60
ttgagtcatt	tgattaaaca	tgtgattatt	taatgaattg	atgagagagt	aggattaaag	120
ttgtattaat	aattagaatt	tgggtgtcaaa	tttaatttga	catttgatct	tttcctatat	180
attgccccat	agagtcagtt	tactcatttt	tatatattcat	agatcaaata	agagaaataa	240
cggatatatta	atccctccaa	caaaaaaaaaa	aaaaaaaaaa	aacggtatat	ttactaaaaa	300
atctaagcca	cgtaggagga	taacatccaa	tccaaccaat	cacaacaatc	ctgatgagat	360
aaccaccttt	aagcccacgc	actctgtggc	acatctacat	tatctaaatc	acacattctt	420
ccacacatct	gagccacaca	aaaaccaatc	cacatcttta	tcacccattc	tataaaaaat	480
cacactttgt	gagtctacac	tttgattccc	ttcaaacaca	tacaaagaga	agagactaat	540
taattaatta	atcatcttga	gagaaaatgg	ccaccgttac	ttccaccacc	gttgctattc	600
catcattcac	aggccttaag	gcaaacgcaa	gcaaagttaa	tgccatagct	aaggttccaa	660
cttcaacttc	tcaattgcca	aggctttgtg	tcagagcttc	cctcaaagac	tttggagttg	720
ctgctgttgc	cactgctgca	agtgcatgtg	tagctagcaa	tgcccttgca	gttgaagtgt	780
tgcttggtgc	tagtgatggg	ggtttggtct	ttgttccaaa	caatttcaca	gtgaacgctg	840
gagacaccat	tacattcaag	aacaatgctg	gttttccctc	caacgttatc	ttcgatgaag	900
acgagattcc	aagcgggggt	gatgctgcaa	tcgaattccc			940

<210> 6

<211> 1368

<212> DNA

<213> Artificial Sequence

<220>

<223> Plastocyanin terminator derived promoter

<400> 6

tcacgcagga	gctgggatgg	ttggacaagt	cactgttaat	taagttaaaa	tgcttcttcg	60
tctcctattt	ataatatggg	ttgttattgt	taattttgtt	cttgtagaag	agcttaatta	120
atcgttgttg	ttatgaaata	ctatttgtat	gagatgaact	ggtgtaatgt	aattcattta	180
cataagtggg	gtcagaatca	gaatgtttcc	tccataacta	actagacatg	aagacctgcc	240
gcgtaacaatt	gtcttatatt	tgaacaacta	aaattgaaca	tcttttgcca	caactttata	300
agtggttaat	atagctcaaa	tatatgggtc	agttcaatag	attaataatg	gaaatatcag	360
ttatcgaaat	tcattaacaa	tcaacttaac	gttatttaact	actaatttta	tatcatcccc	420
tttgataaat	gatagtacac	caattaggaa	ggaaaaatgt	ggtaacacta	atatatgtca	480
gaggatgaac	atctatttca	tattttgatc	ttaacacgac	ttttaaatat	aaaattaaaa	540
caccaatgta	taaattaaat	ttcaaacaat	atattacctc	gaaagtgaag	ttctcattgg	600
aacaattata	ataacattaa	catccttgaa	aacattggcc	gtcgggaattg	gtgtggagct	660
aggaaatgat	ggtttggtat	taagcacctg	aagcccccaa	actttatgta	acaaacaatc	720
aatacaaaac	aatatacagt	gtggctgtgg	tatatgcagc	ctattgcatg	tagatggaaa	780
agaaaaatcc	agggacaagg	agttgtttaa	tttttgccta	gcacgtcatg	cgatccaatt	840
ctctccgtta	tttgagaacg	ggtaacttatc	ttaggtagat	ttttttgtgt	gataaaaaat	900
agtataggcg	agtaggggtt	cagatgctca	ctagagatca	atttaataag	aatttatgag	960
gttatgtcta	cgtgggttag	gcatttggaa	gggtagaacg	ggaaccaagc	actttgccgc	1020
caacgtctgg	gaaagactca	aatccaggta	tagaacttac	tttcccattt	gcattcactt	1080
tgatggggta	cttcataatg	tgcccttgca	atgggtttaa	atcatcatct	gtaaacttct	1140
tccagttgtc	ttcagcaatc	ttgttcacac	tctccacaca	ttccaaactt	tcgggctcct	1200
tgaagcaate	atgtttgggc	cctaagtgct	ctgcccacaa	ggacattcta	taccatata	1260
cctgcaataa	tgtcaattaa	acaagtcctt	gtgaaagtga	taaattggaa	atttcctoca	1320
gccgctaaaa	ggaccagccc	gggcccgtcg	ccacgcgtgc	cctatagt		1368

<210> 7

<211> 27

<212> DNA
<213> Artificial Sequence

<220>
<223> Sense primer NIR promoter-AP2

<400> 7
ttgtcacatc agcacatccg tctttgc 27

<210> 8
<211> 27
<212> DNA
<213> Artificial Sequence.

<220>
<223> Antisense primer NIR promoter AP2

<400> 8
tcgccaagta tcttgtttga gcacttg 27

<210> 9
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Plastocyanin derived sense primer

<400> 9
atgtcttcct tctcagtacg tttctc 27

<210> 10
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Plastocyanin adapted antisense primer

<400> 10
caagttgatg catcaaggtt ggatcctaga 30

<210> 11
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Plastocyanin promoter derived sense primer

<400> 11
aggagcattg agaagatctt cttcagg 27

<210> 12
<211> 27
<212> DNA
<213> Artificial Sequence

<220>

<223> Plastocyanin promoter adapted antisense primer

<400> 12

gctgcatcaa ccccgcttgg aatctcg

27

<210> 13

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Plastocyanin terminator sense primer

<400> 13

gcgttacttt ggatgctaag ggaacct

27

<210> 14

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Plastocyanin terminator antisens primer

<400> 14

tcacgcagga gctggtatgg ttggaca

27

<210> 15

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> NIR derived sense primer

<400> 15

gatctcccta acagtctcaa aagtgt

26

<210> 16

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> NIR derived antisense primer

<400> 16

ggtttctaca ggacgtaaca tttttggaga agagagtgtg tttgg

45

<210> 17

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> GUS ATG primer

<400> 17

atgttacgctc ctgtagaaac c

21

<210> 18

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> NOS terminator primer

<400> 18

gccatgaatt cccgatctag taacatag

28

<210> 19

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> SAC primer

<400> 19

agaagagctc ttgtacattt ggataagtca

30

<210> 20

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Eco primer

<400> 20

agaagaattc gttttcccgga tacttcaact

30

<210> 21

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> GUS SAC primer

<400> 21

agaagagctc gttaaaatgc ttcttcgtct ccta

34

<210> 22

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> GUS Eco primer

<400> 22

agaagaattc tccttcctaa ttggtgtact atca

34